

MOLECULAR EVOLUTION OF EF1 α AND ITS UTILITY FOR RESOLVING
PHYLOGENIES IN THE EUGLENOZOA AND EUGLENIDA

by

SARAH E. JARDELEZA

(Under the Direction of Mark A. Farmer)

ABSTRACT

The gene elongation factor 1 alpha (EF1 α) has had a complicated and somewhat unclear evolutionary history; however, it still remains a useful tool in the reconstruction of organismal phylogeny at varying levels in eukaryotic species. In this study, the relationship is explored between EF1 α and a seemingly-related protein, elongation factor-like (EFL), that sometimes is found to have replaced the function of EF1 α . We found evidence that EF1 α and EFL are ancient paralogs, possibly of the archaeal version of EF1 α , that have also undergone more recent events of horizontal gene transfer (HGT) in certain eukaryotic lineages. Although some members of the euglenozoan (Excavata) lineage, the diplomonids, seem to have completely lost EF1 α the remaining lineages and sister group to the euglenozoans were clarified with both EF1 α and small subunit ribosomal DNA (18S). The discicristates clade was supported with the Heterolobosea found as sister to the Euglenozoa. The recently described Discoba clade was also supported with the Jakobida recovered as sister to the discicristates. The Euglenida lineage was further analyzed with EF1 α , 18S and large subunit ribosomal DNA (28S) to recover supported clades. However, it was determined that within the Euglenida there have been multiple independent gene duplication events of EF1 α , which are useful for the genus and species-level clarification of the

Euglenida organisms, but can obscure the true species history of the group at deeper levels if used inappropriately. EF1 α is still considered useful in phylogenetic reconstruction since it is such a conserved gene, but the study and understanding of the gene within the context of the lineage to be studied must be determined prior to analyses.

INDEX WORDS: elongation factor 1 alpha (EF1 α), elongation factor-like (EFL), 18S, small subunit ribosomal DNA, 28S, large subunit ribosomal DNA, eukaryotes, discicristates, Euglenozoa, Heterolobosea, Discoba, Jakobida, Euglenida, euglenids, Kinetoplastida, kinetoplastids, Diplonemea, diplonemids, tree of life (ToL), phylogenetic analysis, systematics, molecular evolution, gene family, horizontal gene transfer (HGT), ancient paralogy, gene duplication, gene loss

MOLECULAR EVOLUTION OF EF1 α AND ITS UTILITY FOR RESOLVING
PHYLOGENIES IN THE EUGLENOZOA AND EUGLENIDA

by

SARAH E. JARDELEZA

B.A., St. Mary's College of Maryland, 2003

A Dissertation Submitted to the Graduate Faculty of the University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2010

© 2010

Sarah E. Jardeleza

All Rights Reserved

MOLECULAR EVOLUTION OF EF1 α AND ITS UTILITY FOR RESOLVING
PHYLOGENIES IN THE EUGLENOZOA AND EUGLENIDA

by

SARAH E. JARDELEZA

Major Professor: Mark A. Farmer

Committee: James Leebens-Mack
Richard Triemer
Jessica Kissinger
Joseph McHugh
Marshall Darley

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2010

DEDICATION

This work is dedicated to my husband, C.J. Winger. This thesis would not have been completed without his unconditional love and support and for that I will always be grateful. I love you, C.J.!

ACKNOWLEDGEMENTS

I would like to thank the undergraduate lab assistants that helped to support the research in this project: Ellen Fragaszy, Sharanya Ragunath, Linh Nguyen, Elizabeth (Vinson) Schulz, Caitlin Sloan, Stephen Thurston, Brandon Pope, Joy Ishii, Darline Oyemakinwa, and Amy Ishii. I would also like to thank Matt Bennett for his consultation on culture and molecular protocol troubleshooting, Christina Castlejohn's friendship as well as for her support in utilizing several phylogenetic programs and the rCluster at UGA, and the NSF PEET grant for funding my Ph.D.

I also want to thank the members of my committee for their help in training me both in course work and throughout the course of this thesis research to become a better scientist. I especially appreciate the help of Jim Leebens-Mack on Chapter 1 of this thesis and Jessie Kissinger on Chapter 4 of this thesis. I want to thank Rich Triemer for helping to train me to be a better protistologist and Joe McHugh for helping me to become a better systematist. Of course there are no words to describe Marshall Darley and all that he has helped me with to become a better protistologist and instructor. He has always supported me and guided me towards pursuing my love of teaching and I can't thank him enough.

My biggest shout out goes to John Shields who was my substitute advisor for much of my graduate career and offered countless hours of humor, advice, tutelage and coffee! Also, it was an honor to be your best T.A. ever!

None of this work would have been possible if it wasn't for Mark Farmer. I greatly appreciate his support, advising and instruction, especially on how to find balance between work and home. He is a marvelous example of how to be a leader and a mentor.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
CHAPTER	
1 ORIGINS, ANCIENT PARALOGY AND HORIZONTAL GENE	
TRANSFER OF EUKARYOTIC EF1 α AND EFL	1
Abstract	1
Introduction	2
Materials and Methods	4
Results	11
Discussion	16
Conclusion	21
References	23
2 EUGLENOZOAN PHYLOGENIES AND EXCAVATE SISTER	
GROUP ELUCIDATION WITH EF1 α AND 18S	130
Abstract	130
Introduction	131

	Materials and Methods.....	136
	Results.....	141
	Discussion.....	145
	Conclusion	147
	References.....	150
3	EUGLENID MULTIGENE PHYLOGENIES WITH EF1 α , 18S	
	AND 28S.....	217
	Abstract.....	217
	Introduction.....	218
	Materials and Methods.....	222
	Results.....	225
	Discussion.....	229
	Conclusion	233
	References.....	236
4	EF1 α EVOLUTION IN THE EUGLENID LINEAGE.....	295
	Abstract.....	295
	Introduction.....	296
	Materials and Methods.....	299
	Results.....	302
	Discussion.....	306
	Conclusion	310
	References.....	312

CHAPTER 1

ORIGINS, ANCIENT PARALOGY AND HORIZONTAL GENE TRANSFER OF EUKARYOTIC EF1 α AND EFL

Abstract:

Elongation factor 1 alpha (EF1 α) is a protein with an essential role in eukaryotic translation, but a similar and seemingly-related protein, Elongation factor-like (EFL), has been found to replace the “endogenous” EF1 α and perform the same function. There are three possible explanations as to the origins of these two proteins and their presence in eukaryotes across the Tree of Life (ToL). The first is that EFL has replaced the endogenous EF1 α via horizontal gene transfer (HGT) from some outside source or vice versa. The second is that EF1 α and EFL are ancient paralogs that were both present in the most recent common ancestor of all eukaryotes (MRCAe) and have undergone differential loss in various lineages across the ToL. The final option is a combination of the other two in which EF1 α and EFL are ancient paralogs, yet in some lineages show evidence of more recent HGT events as opposed to the differential loss of the original paralogs. Although most recent studies have invoked the first hypothesis to explain the relationship between EF1 α and EFL, we present the first data that provide direct support of the ancient paralogy of EF1 α and EFL in the MRCAe as well as support of recent HGT events using taxonomic sampling across the ToL.

Introduction:

Elongation factor 1 alpha (EF1 α) is an essential gene involved in eukaryotic translation. Its role is to deliver aminoacyl tRNAs to the ribosome during translation. It is also characterized as a nuclear-encoded slowly-evolving protein that has been successfully used in phylogenetic studies to resolve deep-level relationships of organisms (Baldauf 2003; Tanabe et al. 2004; James et al. 2006). However, Keeling and Inagaki (2004) found that in some lineages the endogenous EF1 α was lost and the gene that purportedly replaced the function of EF1 α was a newly discovered and described gene, elongation factor-like (EFL). At the time it was described, EFL was thought to be very limited in its distribution throughout the Eukaryotic Tree of Life (ToL), however, since then EFL has been found in organisms in each of the six main supergroups except Amoebozoa (Adl et al. 2005; Gile, Patron, Keeling 2006; Ribichich, Georg, Gomes 2006; Noble, Rogers, Keeling 2007; Gile, Keeling 2008; Kamikawa, Inagaki, Sako 2008; Marshall et al. 2008; Yoon et al. 2008; Cocquyt et al. 2009; Gile et al. 2009a; Gile et al. 2009b; Sakaguchi et al. 2009).

Unlike previous studies that have examined specific clades and the relationships of EFL and EF1 α in those lineages (Gile, Patron, Keeling 2006; Gile, Keeling 2008; Kamikawa, Inagaki, Sako 2008; Cocquyt et al. 2009; Gile et al. 2009a; Gile et al. 2009b; Sakaguchi et al. 2009), this study encompasses taxa from across the Eukaryotic ToL for both genes in a single analysis, which has not been done since the first description of the gene EFL (Keeling, Inagaki 2004). To utilize as much of the phylogenetic information in the genes as possible, we are using the nucleotide sequence versus the more highly conserved amino acid sequence that other studies have preferentially used when studying these genes, which has also never been done to recover the relationships between EF1 α and EFL. Due to the degeneracy of the amino acid code and the

ancient divergences of major eukaryotic lineages, the third codon position of the genes may be saturated with mutation events that cannot be corrected by parameters of Models of Sequence Evolution (MoSE) in Maximum Likelihood (ML) or Maximum Likelihood Estimation (MLE) analyses, such as gamma correction and general-time-reversible models with six rate categories. We created saturation plots to examine the presence of this phenomenon and performed analyses with the third codon positions removed to detect if there was any difference in clade recovery and whether those differences were affecting deeper nodes or tips in an effort to detect if the third codon positions represented added noise or historical signal.

There are three possible explanations as to the origins of these two proteins and their presence in eukaryotes across the Tree of Life (ToL). The first is that EFL has replaced the endogenous EF1 α via horizontal gene transfer (HGT) from some outside source or vice versa. The second is that EF1 α and EFL are ancient paralogs that were both present in the most recent common ancestor of all eukaryotes (MRCAe) and have undergone differential loss in various lineages across the ToL. The final option is a combination of the previous two in which EF1 α and EFL are ancient paralogs, yet in some lineages there has been more recent HGT events as opposed to the differential loss of the original paralogs.

Although most recent studies have invoked the first hypothesis to explain the relationship between EF1 α and EFL (Kamikawa, Inagaki, Sako 2008; Gile et al. 2009b; Sakaguchi et al. 2009), we present the first data that provide direct support of the ancient paralogy of EF1 α and EFL in the MRCAe by using taxonomic sampling across the ToL. EF1 α and EFL have been supported across the ToL as being mutually exclusive, except for the recent case of *Thalassiosira pseudonana* which appears to be a case of HGT (Kamikawa, Inagaki, Sako 2008). However, we have found mRNA copies of EF1 α and EFL in *Xenopus laevis* (African clawed frog) that might

have had MRCAe as their most recent common ancestor. In phylogenetic analyses, the EF1 α of *X. laevis* clusters with the EF1 α of metazoans, choanomonads and ichthyosporeans (Opisthokonta) and similarly the EFL of *X. laevis* clusters with the EFL of other opisthokonts. A probable origin of the ancestor of the eukaryotic EF1 α /EFL seems to be the archaeal version of EF1 α (aEF1 α). We also found support for recently proposed HGT events such as diatoms receiving their current EFL from foraminiferans (Kamikawa, Inagaki, Sako 2008). These data support the third explanation, in which the current distribution of EF1 α and EFL across the eukaryotic ToL is the result of a combination of differential loss of the ancient paralogs as well as more recent HGT events.

Materials and Methods:

Culture information and nucleic acids extractions

All taxa with sources and accession numbers used in this study are shown in Table 1.1. Taxa were either obtained from culture collections or from colleagues (Triemer et al. 2006). All isolates were grown in either modified AF-6 (Watanabe, Hiroki 1997), saltwater AF-6 for *Eutreptiella marina* and *Eutreptiella pomquetensis*, hay infusion for *Entosiphon* sp., *Peranema trichophorum*, *Ploeotia oblonga* or CCAP ASWP (Culture Collection of Algae and Protozoa Artificial Sea Water for Protozoa) for *Ploeotia costatum*. Cultures were maintained at 20-25°C, except *Eutreptiella pomquetensis* (4°C) and *Eutreptiella marina* (15°C), under conditions of a 12:12 light:dark cycle and approximately 30 $\mu\text{mol photons m}^{-2}\text{s}^{-1}$ from cool white fluorescent tubes.

Sequence retrieval, nucleotide isolation, amplification, and sequencing

Species from Archaea (5), Amoebozoa (4), Archaeplastida (24), Chromalveolata (22), Excavata (92), Opisthokonta (29), Rhizaria (5) and several other sequences from either an

environmentally sampled unidentified species (1) or historically originating from endosymbiosed organisms (3) were tested for the presence of EF1 α and EFL and/or obtained with PCR or searching EST and genomic databases (JGI, TBestDB, TriTrypDB v.1.1) or GenBank release 172 (Table 1.1). New sequences were deposited in GenBank with accession numbers XXXX-XXXX: euglenid (46), kinetoplastid (1), heterolobosean (1). Sequences that had been used in prior studies, but were not given GenBank accession numbers specific to that gene were also deposited with accession numbers XXXX-XXXX (this includes 4 Archaea, *S. obliquus* (chlorophyceae), *O. tauri* (prasinophyte), *P. marinus* (dinoflagellate), *E. huxleyi* (haptophyte), *A. macrogynus* and *S. punctatus* (chytrids).

Total genomic DNA (gDNA) was isolated from uniprotistan cultures with either DNeasy Plant Mini Kit (Qiagen Co., Valencia, CA, USA; Cat# 69104) or MoBio Ultraclean™ Microbial DNA Isolation Kit (MoBio Laboratories, Inc., Carlsbad, CA; Cat# 12224-250). Total RNA was extracted from uniprotistan cultures with RNeasy Plant Mini Kit (Qiagen; Cat# 74903) and then converted to cDNA using Superscript II Reverse Transcriptase (Invitrogen Corp., Carlsbad, CA; Cat# 18064-014) with Oligo(dT) primer (Invitrogen, Cat# 18418-012). The gDNA or cDNA template was used with the following PCR master mix to amplify EF1 α sequences using the MJ Research or Bio-Rad PTC-200 (Bio-Rad Inc., Hercules, CA, USA): 12.4 μ L of molecular-grade water, 2.5 μ L of 10x Bioline Buffer, 1 μ L of DMSO, 5 μ M MgCl₂, 4 mM dNTPs, 0.6 μ M of each primer (forward and reverse), 1 μ L of DNA (~25-300 ng) and 0.1 μ L of Biolase DNA Polymerase (Bioline, Cat# BIO-21043). Thirty-five cycles with annealing temperatures of either 47 or 52°C and gradually increasing extension times for each cycle were required for the EF1 α -extended PCR program using EUG EF1a 1F, EUG EF1a 1R, EUG EF1a 1Rgg, and EUG EF1a 2R (extended PCR details in Table 1.2, primer details in Table 1.3). The longer extension times

are required to allow the polymerase to complete the product since its enzymatic activity decreases as the length of the program increases. When the primer pairs GG_EF1a_F3 and Bal_10R or 1F_FL and Bal_7R2 were used, then the PCR program was more standard with annealing temperatures of either 50.4 or 37°C, respectively (Table 1.2).

The PCR products were sized on 1% agarose gels and then purified using the MinElute Gel Extraction Kit (Qiagen Co.; Cat #28606) according to the manufacturer's protocol. The purified template was either sequenced directly with the original primers on an ABI 3730 Capillary Electrophoresis Genetic Analyzer using the ABI BigDye Terminator v3.1 Cycle Sequencing (Perkin-Elmer Applied Biosystems, Foster City, CA, USA) or cloned into bacterial vector using either TOPO TA Cloning® Kit with pCR® 2.1-TOPO® Vector (Invitrogen; Cat# K451020) or pGEM® -T Easy Vector System (Promega Corp., Madison, WI; Cat# A1360). If the latter, then 1-10 clones were extracted from the plasmids using Plasmid Mini Kit (Qiagen Co.; Cat# 12123) and then sequenced using m13r and T7 primers associated with the vector on an ABI 3730xl DNA Analyzer.

Sequence Alignment

Nucleotide sequences were aligned with MUSCLE (Edgar 2004) and then edited by eye against closely related organisms' EF1 α or EFL sequences based on amino acid conceptual translations in Geneious Pro v3.8.5 (Biomatters Ltd.; available through www.geneious.com) (Drummond et al. 2009) and previously published EF1 α and EFL alignments (Baldauf, Palmer 1993; Baldauf, Doolittle 1997; Keeling, Inagaki 2004; Fiore-Donno et al. 2005; Gile, Patron, Keeling 2006; Gile, Keeling 2008). Those gapped alignments (Fig. 1.1) were either analyzed directly or were run through GBlocks v0.91b (Castresana 2000; Talavera, Castresana 2007) on the online server (http://molevol.cmima.csic.es/castresana/Gblocks_server.html) to create

reproducible alignments with unsupported positions eliminated. The GBlocks parameters selected allowed as much retention of the original alignment as possible and were as follows: codon input data, allowed for smaller blocks, allowed for gap positions within final blocks and allowed for less strict flanking positions. An intermediate alignment was also created that maintained some known phylogenetically informative indels that would have been excised by GBlocks. These alignments will be referred to as Muscle output (Mo), Muscle output followed by GBlocks output (MoGo) and Muscle output with limited cuts from GBlocks output (MolcGo), respectively. Then, third codon nucleotides were removed from all three alignments to eliminate any homoplastic noise associated with that position to create three new alignments: Mo3rd, MoGo3rd and MolcGo3rd, respectively. Saturation plots were made using the empirical alignment data or the RAxML estimated parameters (see details below for RAxML analyses) to compare each original alignment with all nucleotides included vs. its respective alignment with the third codon positions removed based on pairwise likelihood distances calculated in PAUP* (Fig. 1.2) (Swofford 1993). All details for the alignments are provided in Table 1.4.

Amino acid sequences were processed similarly to their nucleotide counterparts, resulting in three new alignments: aaMo, aaMolcGo and aaMoGo (Fig. 1.3). The GBlocks server parameters were the same as mentioned previously, except the input data was selected as protein.

Phylogenetic analysis

All-nucleotide alignments with 194 operational taxonomic units (OTUs) across the ToL from archaeal EF1 α (aEF1 α), and eukaryotic EF1 α (EF1 α) and EFL were processed with Maximum Parsimony (MP), Maximum likelihood (ML), and Maximum likelihood estimation using Bayesian algorithms (MLE) (Table 1.5). Alignments were analyzed for parsimony informative characters (Table 1.4) and prepared for MP analysis by removing parsimony

uninformative characters in Winclada v1.0000 (Nixon 1999-2002). MP (single run) and MP bootstrapping (1000 bootstrap runs) were performed using a new technology search with sectorial search, ratchet, drift and tree fusing selected in TNT v1.1 (Goloboff, Farris, Nixon 2004; Goloboff, Farris, Nixon 2008) on the Phylogeny.fr online server (http://phylogeny.lirmm.fr/phylo_cgi/index.cgi) (Dereeper et al. 2008). ML was performed using both RAxML v7.0.4 with simultaneous autodetect rapid bootstrapping (Stamatakis 2006; Stamatakis, Hoover, Rougemont 2008) and GARLI v0.951 (Zwickl 2006) analyses and 100 bootstrap replicates on the CIPRES server (Miller et al. 2009). MLE was also performed with MrBayes v3.1.2 (Huelsenbeck, Ronquist 2001; Ronquist, Huelsenbeck 2003) on either the CIPRES server or the University of Georgia (UGA) rCluster. Parameters for all analysis types are available in Table 1.5. For the ML and MLE analyses, the model used was the optimal model agreed upon by both the hLRT and AIC indices from MODELTEST v3.7 (Posada, Crandall 1998; Posada, Buckley 2004). The best model for all alignments was GTR+I+ Γ . To ensure convergence upon the correct topology MLE consensus using the nucleotide alignments, four separate runs of thirty million generations each were performed with trees sampled every thousand generations. Of the thirty thousand trees created from each run a conservative fifteen million trees were discarded as burn-in. The resulting fifteen thousand trees from each of the four runs were then combined for a total of sixty thousand trees in PAUP* in a majority rules consensus tree to determine relationships between clades as well as posterior probability (pp) support of the MLE topology. Trees were rooted using the archaeal sequences of EF1 α because they have been shown to have a close affinity to eukaryotic EF1 α and EFL (Keeling, Inagaki 2004; Jardeleza, unpub. data).

As a further verification that nucleotide analyses' findings were appropriate, given these data, the more conserved amino acid alignments were also analyzed. They were processed for phylogenetic analysis as described for their nucleotide counterparts, with a few exceptions. The appropriate MoSE was determined with ProtTest v1.4 (Abascal, Zardoya, Posada 2005) for each alignment (Table 1.4). The best model for all amino acid alignments was RTREV+I+ Γ +F. Only RAxML and MrBayes were used to analyze the data sets, with parameters specified in Table 1.6. MrBayes analyses were performed on the new Abe teragrid on the CIPRES server (Miller et al. 2009), which was not originally available for the large nucleotide analyses. The MrBayes amino acid analyses used six runs of six million generations with three million generations discarded as burn-in (Table 1.6). Also unlike the EF1 α single-gene alignments, these runs were sampled every hundred generations resulting in a total of 180 thousand trees in the final consensus tree for each alignment, which was used to determine relationships of clades as well as the pp value for each node that was recovered. Rooting remained the same as in nucleotide analyses.

Clade Recovery

All alignments were considered using all analysis types and recovery of expected monophyletic clades based on previous literature were summarized (Table 1.7). A Phylocode-like hierarchical scheme is employed in Table 1.7 to help visualize relative depth in the Eukaryotic branch of the ToL; however, the rank status is not comparable between groups. Clades are represented not only with “supergroups” and “groups” names listed in Table 1.1, but more recent clades are also shown that represent larger classification types, such as the bikonts, the unikonts (Burki, Pawlowski 2006; Burki, Shalchian-Tabrizi, Pawlowski 2008) and the SAR (Stramenopila, Alveolata and Rhizaria) clades (Hackett et al. 2007). The currently un-named but historically robustly supported clade consisting of the Choanomonada, Mesomycetozoa and

Metazoa group was termed the CMM clade for the purpose of this study. The Archaeplastida EFL clade is assumed to include the copy targeted to the periplastid compartment (ppc) of EFL from Chlorarachniophytes since that copy is associated with the reduced endosymbiont that was originally supported to be a green alga (Gilson et al. 2006; Gile, Keeling 2008). The CMM EFL clade includes the uncultured eukaryote species EFL sequence that consistently branched with that clade in a previous study (Keeling, Inagaki 2004). *Oryza sativa* EFL is not taken into consideration as a charophyte/Streptophytina clade-member because it is deemed an EFL sequence of some unknown green algal contaminant since the sequence is not recovered in the *O. sativa* genome sequence databases (Keeling, Inagaki 2004). The total number of recovered clades out of the ninety clades we expected to be able to recover, if the gene trees match the species tree, was tabulated for each alignment type in order to analyze which was best at clade recovery and therefore the optimal alignment for this study to be used in subsequent tree topology tests. Relative depth of clades recovered, with deeper clades preferred, and possible signal saturation at the third codon position were also taken into account when determining the optimal alignment.

Tree Topology Tests

Completely bifurcating single best trees from all runs of all analyses using the optimal alignment as decided by the aforementioned criteria were analyzed for which is the “best” tree or trees. To perform an approximately unbiased (AU) tree topology test, all possible trees and the MolcGo3rd alignment were used to estimate per-site likelihoods (PSLs) for each tree with RAxML v7.0.4 on the UGA rCluster. The resulting PSLs were then analyzed by CONSEL v1.0 (Shimodaira, Hasegawa 2001) to rank the trees by their AU score. Trees that could not be statistically rejected ($p\text{-value} > 0.05$) were analyzed and conclusions about gene and species

relationships as well as phylogenetic analysis recovery type (e.g., MP, ML, MLE) were made from them. We also created a general tree that was based on clades that we assumed to be recovered (Table 1.1) and changed the rooting of the gene tree to test the alternative hypotheses of the ancestral origin of the three GTPase gene family genes in this study with either aEF1 α basal to the clade of EFL+EF1 α , EF1 α basal to the clade of aEF1 α +EFL or EFL basal to the clade of aEF1 α +EF1 α . We also used distance methods and used midpoint rooting to help clarify this issue.

Results:

Effects of Alignments and Analysis Type and their Impact on Clade Recovery

There was variation in the resulting topologies based on which nucleotide alignment and analysis type were used; however, many organismal clades that were expected to be found were recovered despite those differences (Fig. 1.4 for the summary gene tree, Fig. 1.5 for the detailed EF1 α portion of the gene tree, and Fig. 1.6 for the detailed EFL portion of the gene tree). Thirty-four of the ninety (~38%) expected clades were recovered with support across all six alignments. When those are further broken down, ~67% (2/3) aEF1 α , ~35% (17/48) EF1 α and ~38% (15/39) EFL clades are recovered. For some of these genes, this percentage of recovered clades increases when the third codon position is excluded; 100% (3/3) aEF1 α and ~44% (21/48) EF1 α clades recovered using Mo3rd, MolcGo3rd and MoGo3rd.

Although the pairwise distance plots exhibited minor amounts of saturation of the third codon position, based on R^2 values less than one, it seems that there are differences between the alignments that contain all residues versus those that have the third codon position removed in recovering certain clades (Fig. 1.2). In particular, the alignments with third codon positions removed (Mo3rd, MolcGo3rd and MoGo3rd) always recover certain basal clades with support

including Opisthokonta, Euglenozoa, Euglenida and Jakobida in EF1 α whereas those clades are never recovered in any alignment including all nucleotides (Mo, MolcGo, MoGo). There are never any clades recovered in all three all-nucleotide alignments compared to the alignments with third codon positions removed. However, there are some relatively recently diverged clades recovered by one or more all-nucleotide alignments that are not recovered in any of the alignments with third codon positions removed, (e.g., Basidiomycota and Metazoa in EF1 α and Coscodiscophytina and Blastocladiaceae in EFL).

There seemed to be a large number of data points within the pairwise distance plot that had Y-coordinates (distances generated using all-nucleotide alignments) that were much higher than expected when compared to the trend line, which could be related to LBA issues or specific sequences were inconsistent in their placement in the gene tree depending upon which alignment was used to create the topology. These points were analyzed and evaluated. Most of the sequences involved in these high pairwise Y-coordinate values were EFL sequences.

Diplonemids, haptophytes, and some Chloroplastida sequences were consistently recovered with high values in all three saturation plot comparisons between the all-nucleotide alignments and their counterparts with the third codon positions removed. The genus members of *Diplonema* are recovered in MolcGo3rd using GARLI, whereas they are not recovered with that analysis type in MolcGo. Haptophytes are always recovered in nucleotide analyses by RAXML, GARLI and MrBayes, except in MoGo3rd, where only RAXML recovers the clade. The Chloroplastida OTUs that are involved in these high values also jump around in different places on the tree, including *Tetraselmis tetrathele* and *Micromonas pusilla* and they do not tend to be located within the main cluster of Chloroplastida OTUs. The only eukaryotic EF1 α sequences involved in high Y-axis pairwise distance values were *Sawyeria marylandensis* and *Plasmodium*

falciparum, both known to have problems with LBA, as well as *Lepocinclis spirogyroides* clone 38 which has a short sequence. Two of the five aEF1 α sequences, *Methanocaldococcus jannaschii* and *Sulfolobus solfataricus*, were also above the trend line, but generally at lower Y-axis values than the aforementioned sequences.

The amino acid analyses offered another verification of the effectiveness of nucleotide MoSE correction and any evidence of possible saturation or hyper-mutation in the third codon positions of the data set sequences (Fig. 1.7 for the summary gene tree, Fig. 1.8 for the detailed EF1 α portion of the gene tree, and Fig. 1.9 for the detailed EFL portion of the gene tree). For the most part, there is a great deal of overlap in the clades recovered by the amino acid and nucleotide analyses. However, there are some instances where amino acid analyses recover groups that nucleotide analyses didn't recover at all or with less analysis types supporting, such as: Rhizaria in EFL, as well as Stramenopila and Basidiomycota in EF1 α . It seems that amino acid lacks the deep-level resolving power that the nucleotide analyses provide since the former do not recover the following clades (even or with less shared analysis types supporting the relationships) that were recovered in the latter: unikonts, Amoebozoa, Metazoa, Ciliophora, "Group 1" excavates, *Malawimonas*, Oxymonadida, Euglenozoa, Trypanosomatida, Vahlkampfiidae, Jakobida, and Histioglossidae in EF1 α as well as Blastocladales, Blastocladiaceae, Zygomycota, Bacillariophytina, Coscinodiscophytina, Dinophyta, and *Diplonema* in EFL.

Archaeal Origin of Ancestral Paralog

When we re-rooted the tree using EF1 α or EFL, we could not statistically reject those alternative topologies. However, the tree rooted with aEF1 α was found to be supported as the best topology based upon the highest AU score for the majority of the alignments. The distance midpoint rooting preferred a closer relationship between aEF1 α and EF1 α , with EFL sister to

them. EFL has a very long branch leading to it and that skews the true rooting history of the genes.

No known Archaea have EFL, yet EFL and EF1 α are scattered among the Eukaryotes across the ToL. It is the most parsimonious explanation that the MRCAe received its original version of EF1 α from an EF1 α and that there was an ancestral gene duplication event of the aEF1 α that led to the current endogenous copies of the ancient paralogs EF1 α and EFL. The alternate option is that the last universal common ancestor of all living organisms (LUCA) had both EFL and a version of EF1 α and that the EF1 α diverged into its present forms, aEF1 α and EF1 α , which would then have to be ancient paralogs or so extremely diverged in the archaeal and eukaryotic lineages that they can still be distinguished as separate gene types in the gene family analyses.

Conflict Between Species Trees vs. Gene Trees

Although some deep-level monophyletic clades are recovered in the nucleotide analyses in the EF1 α portion of the gene tree, both EF1 α and EFL generally lack resolving power within the main supergroups (Fig. 1.4). The Amoebozoa and Opisthokonta are recovered as monophyletic based on the organisms in those groups that still retain EF1 α , as well as the larger megagroup known as the unikonts that combines those two clades. However, the Archaeplastida, Chromalveolata, Excavata and Rhizaria clades are never recovered as monophyletic in either the EF1 α or EFL portions of the gene tree. Unfortunately, like Amoebozoa which only retains EF1 α , the Rhizarians only retain EFL. The ancestor leading to all Rhizarians seems to have lost its endogenous EF1 α . Although the two clades of Rhizarians sampled in this study, Chlorarachniophyta and Foraminifera, are consistently recovered, they are never recovered as a monophyletic clade together in the nucleotide analyses, although they are

recovered in the amino acid analyses (Fig. 1.7). They are always found as basal in the EFL portion of the gene tree and may represent the EFL-type of the MRCAe (Figs. 1.4 and 1.7). Similarly, the “Group 1” excavates are always recovered at the base of the nucleotide EF1 α gene branch (Figs. 1.4 and 1.5B). The Excavata are believed to be representative of the MRCAe (Baldauf 2003; Burki, Pawlowski 2006; Burki, Shalchian-Tabrizi, Pawlowski 2008), so this too is supportive of that hypothesis. Many other clades of varying depth in the Eukaryotic ToL are still recovered in both the EFL and EF1 α portions of the gene tree.

Co-occurring EF1 α and EFL

In this study we have three organisms that are represented by both EF1 α and EFL, however, the histories of those genes and their organisms are different. *Oryza sativa* has been shown in previous literature to have EF1 α only. The EFL *O. sativa* sequence is more similar to green algae in BLAST searches, and it is never recovered in BLAST searches in any of the *O. sativa* completed genome project databases. Therefore, it has been posited that the *O. sativa* EFL is most likely a green algal contaminant (Keeling, Inagaki 2004, this study).

Thalassiosira pseudonana is a diatom that has both EFL and EF1 α that are detectable by RT-PCR, however, the very diminished amount of transcript detectable for the EF1 α indicates that it is no longer fully functional in the role of the primary elongation factor (Kamikawa, Inagaki, Sako 2008). This same study hypothesized that all diatoms appear to have had a HGT event of EFL from a foraminiferan donor that seemingly replaced their endogenous elongation factor. This is supported in our findings since the diatoms with EFL are always sister or closely related to the Foraminifera as opposed to other more chromalveolate organisms, such as the Cryptophyteae, Haptophyta or Dinzoa. Therefore, the EFL in *T. pseudonana* is not the endogenous EFL, but a replaced version of the gene, although the EF1 α copy does appear to be

endogenous since it is sometimes recovered with other stramenopiles in our findings (Figs. 1.5B and 1.8B).

However, *Xenopus laevis* (African clawed frog) does appear to maintain functional copies of endogenous EF1 α and EFL, although the levels, tissues and life cycle stages at which the genes are expressed have not been studied. The gene sequences are both from mRNA and are found in BLAST searches to the *X. laevis* completed genome. EFL was present in at least the tadpole stage of the organism, whereas EF1 α was expressed in at least the oocytes. Also, the location in both the nucleotide and amino acid gene trees in this study place *X. laevis* EFL with other members of the CMM clade, nested between ichthyosporeans and a choanomonad, *Monosiga brevicollis* and *X. laevis* EF1 α is sister to the duckbilled platypus within the Metazoan clade of the CMM clade in the Opisthokonta clade of EF1 α (Figs. 1.6 and 1.9, and Figs. 1.5B and 1.8, respectively).

Discussion:

Effect of Third Codon Positions on Clade Recovery

Using the nucleotide sequence of protein-coding genes effectively provides three times the number of characters that can be utilized for resolving phylogenetic relationships, however, the third codon position is known to be problematic for its homoplastic signal when studying deep-level relationships due to the degeneracy of the amino acid code. The EF1 α and EFL gene relationship is certainly deep if both were present in the MRCAe. However, when we check the saturation plots of the alignments containing various amounts of sequence removal due to indels in either lineages, entire genes or just an abundance of incomplete sequences and compare those to their respective alignments with the third codon positions culled, one does not see a great deal of signal saturation for those sites (Fig. 1.2). Instead, we notice outlier organisms that are either

showing more or less signal saturation than is expected based on the model of sequence evolution given the data sets. Since those outlier organisms tend to be members of a long branching gene, EFL, or long-branching sequences, *S. marylandensis* and *P. falciparum*, it seems that the hyper-mutation of the third codon position could be affecting the placement of these sequences between the analyses using different alignment types. Since *Diplonema* monophyly in EFL is recovered by GARLI using MolcGo3rd, but not MolcGo, it seems that this lineage has noise in its third codon position that is confounding the historical relationships of the organisms (Table 1.7 and Fig. 1.6). On the other hand, Haptophyta monophyly in EFL is recovered with less analysis types using the MoGo3rd alignment versus the MoGo alignment, which supported that in this instance the third codon position had signal that was lineage specific that aided in the phylogenetic recovery of this clade when it was included (Table 1.7).

However, when the resulting topologies are examined further distinctions between alignments with all the nucleotides versus those with third codon positions removed become apparent (Table 1.7). Since deeper-level clades such as Opisthokonta, Euglenozoa, Euglenida and Jakobida are recovered in all three nucleotide alignments with the third codon positions removed and are never recovered in the alignments with all nucleotides we conclude that there is saturation in the third codon position for ancient divergences. Conversely, when all-nucleotide alignments are used, the resulting topologies sometimes recover more recently diverged clades such as Basidiomycota, Metazoa, Coscinodiscophytina and Blastocladiaceae that are never recovered when the third codon positions are removed. Both of these findings give indirect evidence that there is undetected homoplastic noise in the third codon position that is not recovered in the saturation plots alone.

Finally, when the third codon positions are excluded, there is greater improvement in clade recovery, especially in the deeper nodes, for EF1 α ; whereas EFL shows no improved clade recovery. On the contrary, EFL seems to recover extra clades when all nucleotides are present, but only tip-level clades. This seems to indicate more recent activity of actively mutating genes, such as the Coscinodiscophytina, which are diatoms with EFL that has been horizontally transferred recently in terms of this study (Fig. 1.6). The long branch to EFL and the reduced clade recovery, especially at deep levels, lends to the view that EFL is a gene that has undergone recent HGT events in several lineages across the ToL (Figs. 1.6 and 1.9).

The amino acid alignments have the benefit of having more character states available for each position, which hypothetically increases the ability of that data set to recover deep-level relationships compared to the nucleotide data sets. Although most clades were recovered with both nucleotide and amino acid analyses, there were clear differences between the two that actually support that the opposite is true and that nucleotide analyses of protein-coding sequences had better resolving power at all levels of the phylogeny as compared to the amino acid analyses (Figs. 1.4 and 1.7). The decreased accuracy of amino acid MoSE to correct for the unseen evolution of the sequences over their evolutionary histories as compared to the well-developed nucleotide MoSE may be more essential to recovering clades than the increased amount of character states that the amino acid data sets provide. It also seems LBA affects amino acid analyses more than their nucleotide counterparts since lineages with known LBA issues are not recovered using the amino acid data sets as opposed to the nucleotide data sets (e.g., Oxymonadida, “Group 1” excavates, *Malawimonas*, Trypanosomatida, Vahlkamfiidae, etc.). Also, some of the lineages that are involved in the high Y-coordinate pairwise distance values in the nucleotide analyses are not recovered in the same topology positions in the amino

acid analyses, such as *Diplonema*, which is never recovered with any analysis type in the amino acid analyses. Similarly, the *T. tetrahele* EFL sequence breaks up the Zygomycota clade in the amino acid analyses rather than being more closely related to the Chloroplastida; which is a relationship that is never found in any of the nucleotide analyses (Figs. 1.6 and 1.9).

Origins and Ancient Paralogy of Eukaryotic EF1 α and EFL

Within each gene either HGT, rampant differential gene loss of ancient paralogs or both mechanisms must be invoked to make sense of the gene trees not reflecting the species tree of EF1 α and EFL. Recent HGT is supported in certain instances (diatom EFL from foraminiferans), however, gene loss and ancient HGT are very difficult to verify or falsify. Indirect evidence of the ancient paralogy of the two genes is their representation in lineages from across the Eukaryotic ToL. This could also be explained, in a very unparsimonious way, by massive amounts of HGT from mostly unknown sources of one gene or the other. Although ancient paralogy of these two genes has been hypothesized since the discovery of EFL (Keeling, Inagaki 2004), this study provides the first direct evidence of ancient paralogy of EF1 α and EFL since *X. laevis* seemingly retains and expresses endogenous EF1 α and EFL.

Keeling and Inagaki (2004) first studied the GTPase gene family including aEF1 α , EF1 α , EFL, HBS1 (heat shock protein 70, subfamily B suppressor 1) and eRF-3 (eukaryotic releasing factor-3). They found that EFL, HBS1 and eRF-3 formed a single clade, but there was a polytomy between that clade and EF1 α and aEF1 α , although the tree was rooted to aEF1 α . In our study we also rooted our trees using aEF1 α , but consistently found that there was a clear sister-relationship between EF1 α and EFL, which supports their status as ancient eukaryotic paralogs originating from aEF1 α . Although we could not reject the alternative rooting tree topologies, the tree rooted with aEF1 α was found to be supported as the best topology based

upon the highest AU score for the majority of the alignments. Statistically speaking, there should not be any difference in AU score for these alternative topologies since rooting changes should not affect the AU for the overall tree and it is unclear why there are differences. This is also supported by the current hypothesis that eukaryotes are derived from a fusion of archaeal and bacterial genetic information (Yutin et al. 2008), such that the aEF1 α could be the potential ancestor of EF1 α and EFL paralogs. Even though the distance midpoint rooting resulted in a closer relationship between aEF1 α and EF1 α , with EFL sister to them, however, this is to be expected based on the long branch and fast site evolution that occurs in the branch leading to the EFL clade as well as the branches within the lineage itself.

The role of Molecular Evolution in Unraveling the History of EFL and EF1 α

The supergroups may not be recovered due to several underlying phylogenetic factors including, but not limited to: lack of signal to reconstruct the true evolutionary history of the genes, inability of the MoSE to accurately correct for the evolution of the sequence(s), homoplasy to confound the history of the genes, long branch attraction artifacts (LBA) of certain species or entire lineages (e.g., EFL), HGT from donors outside of a lineage's original supergroup placement, or differential loss of ancient paralogs (EF1 α or EFL). The low support of branches across the backbone of the EF1 α portion of the gene tree lends support to the explanation of lack of signal and/or increased homoplasy to reconstruct the evolutionary history of that gene. In the case of the ciliates and apicomplexans (Alveolata) in the EF1 α portion of the gene tree, these lineages are fairly fast-evolving (as has been found previously in other genetic data) and although they sometimes branch close to one another in the tree, they are never recovered as the clade Alveolata, and therefore the larger supergroup Chromalveolata can never be recovered. Also, it seems that EFL sequences are allowed relaxed selective pressure and

therefore evolve much more rapidly as compared to their EF1 α counterparts in closely related lineages. Therefore, although we can recover Dinoflagellata, Cryptophyceae and Haptophyta, they are never recovered together into the supergroup they are part of, the Chromalveolata (Figs. 1.4 and 1.6). HGT is also supported when we recover diatoms branching close to or often sister to Foraminifera organisms on the EFL portion of the gene tree (Fig. 1.6). The diatoms share an indel idiosyncratic to only the Foraminifera when EFL is studied across the Eukaryotic ToL (Kamikawa, Inagaki, Sako 2008), lending support to a foraminiferan being the donor to the ancestor that gave rise to all diatoms. However, we cannot discount the differential loss of ancient paralogs when we account for the lineages that are represented in both the EFL and EF1 α portions of the gene tree in monophyletic clades (e.g., Rhodophyceae (Archaeplastida), chytrids, zygomycetes, choanomonads, ichthyosporeans, and metazoans (Opisthokonta), as well as Kinetoplastida (Excavata); Figs. 1.4 and 1.7).

Conclusion:

The gene history of the two genes is seemingly obscured by avenues of molecular evolution, such as LBA and HGT, which have essentially been re-writing the ancient answers present only in the MRCAe, however, direct evidence of the ancient paralogy of the two genes can still be obtained in an extant organism. This is the first study that examines the origin of eukaryotic EF1 α and EFL since the original discovery of EFL (Keeling, Inagaki 2004). This study provides an even larger scope and deeper view of the relationships within and between the genes and their relationship with archaeal EF1 α because of the vast amount of sequence data acquired from genome, EST and PCR projects to give a clearer picture of the history of eukaryotic EF1 α and EFL. Seemingly endogenous and expressed copies of both EF1 α and EFL have been found in *X. laevis*, where the only other evidence of true co-occurrence is found in an

organism where at least one of the genes is supported as originating through HGT (Kamikawa, Inagaki, Sako 2008). This coupled with the sister relationship of the two genes rooted with archaeal EF1 α supports that current eukaryotic EF1 α and EFL are ancient paralogs. The current gene trees do not reflect the current species trees when either nucleotide or amino acid data sets were analyzed and this is most likely due to differential loss of the ancient paralogs and HGT, as well as the limitations of amino acid MoSE. As more organisms are studied and more sequences are determined, different relationships may be elucidated; however, some instances may be so ancient that the true history may be unrecoverable in such a well-conserved gene family.

References:

- Abascal, F, R Zardoya, D Posada. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21:2104-2105.
- Adl, SM, AGB Simpson, MA Farmer, et al. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology* 52:399-451.
- Baldauf, SL. 2003. The deep roots of eukaryotes. *Science* 300:1703-1706.
- Baldauf, SL, WF Doolittle. 1997. Origin and evolution of the slime molds (Mycetozoa). *Proceedings of the National Academy of Sciences of the United States of America* 94:12007-12012.
- Baldauf, SL, JD Palmer. 1993. Animals and Fungi Are Each Others Closest Relatives - Congruent Evidence from Multiple Proteins. *Proceedings of the National Academy of Sciences of the United States of America* 90:11558-11562.
- Burki, F, J Pawlowski. 2006. Monophyly of Rhizaria and multigene phylogeny of unicellular bikonts. *Molecular Biology and Evolution* 23:1922-1930.
- Burki, F, K Shalchian-Tabrizi, J Pawlowski. 2008. Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. *Biology Letters* 4:366-369.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17:540-552.
- Cocquyt, E, H Verbruggen, F Leliaert, FW Zechman, K Sabbe, O De Clerck. 2009. Gain and loss of elongation factor genes in green algae. *Bmc Evolutionary Biology* 9.
- Dereeper, A, V Guignon, G Blanc, et al. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research* 36:W465-W469.
- Drummond, AJ, B Ashton, M Cheung, J Heled, M Kearse, R Moir, S Stones-Havas, T Thierer, A Wilson. 2009. Geneious: Biomatters Ltd.
- Edgar, RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792-1797.
- Fiore-Donno, AM, CD Berney, J Pawlowski, SL Baldauf. 2005. Higher-order phylogeny of plasmodial slime molds (Myxogastria) based on elongation factor 1-A and small subunit rRNA gene sequences. *Journal of Eukaryotic Microbiology* 52:201-210.
- Gile, GH, D Faktorová, CA Castlejohn, G Burger, BF Lang, MA Farmer, J Lukeš, PJ Keeling. 2009a. Distribution and Phylogeny of EFL and EF-1 α in Euglenozoa Suggest Ancestral Co-Occurrence Followed by Differential Loss. *PLoS ONE* 4.

- Gile, GH, PJ Keeling. 2008. Nucleus-encoded periplastid-targeted EFL in chlorarachniophytes. *Molecular Biology and Evolution* 25:1967-1977.
- Gile, GH, PM Novis, DS Cragg, GC Zuccarello, PJ Keeling. 2009b. The Distribution of Elongation Factor-1 Alpha (EF-1 alpha), Elongation Factor-Like (EFL), and a Non-Canonical Genetic Code in the Ulvophyceae: Discrete Genetic Characters Support a Consistent Phylogenetic Framework. *Journal of Eukaryotic Microbiology* 56:367-372.
- Gile, GH, NJ Patron, PJ Keeling. 2006. EFL GTPase in cryptomonads and the distribution of EFL and EF-1 alpha in chromalveolates. *Protist* 157:435-444.
- Gilson, PR, V Su, CH Slamovits, ME Reith, PJ Keeling, GI McFadden. 2006. Complete nucleotide sequence of the chlorarachniophyte nucleomorph: Nature's smallest nucleus. *Proceedings of the National Academy of Sciences of the United States of America* 103:9566-9571.
- Goloboff, PA, JS Farris, KC Nixon. 2004. Tnt. *Cladistics-the International Journal of the Willi Hennig Society* 20:84-84.
- Goloboff, PA, JS Farris, KC Nixon. 2008. TNT, a free program for phylogenetic analysis. *Cladistics-the International Journal of the Willi Hennig Society* 24:774-786.
- Hackett, JD, HS Yoon, S Li, A Reyes-Prieto, SE Rummele, D Bhattacharya. 2007. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of Rhizaria with Chromalveolates. *Molecular Biology and Evolution* 24:1702-1713.
- Huelsenbeck, JP, F Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.
- James, TY, F Kauff, CL Schoch, et al. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443:818-822.
- Kamikawa, R, Y Inagaki, Y Sako. 2008. Direct phylogenetic evidence for lateral transfer of elongation factor-like gene. *Proceedings of the National Academy of Sciences of the United States of America* 105:6965-6969.
- Keeling, PJ, Y Inagaki. 2004. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1 alpha. *Proceedings of the National Academy of Sciences of the United States of America* 101:15380-15385.
- Marshall, WL, G Celio, DJ McLaughlin, ML Berbee. 2008. Multiple isolations of a culturable, motile ichthyosporean (Mesomycetozoa, Opisthokonta), *Creolimax fragrantissima* n. gen., n. sp., from marine invertebrate digestive tracts. *Protist* 159:415-433.
- Miller, MA, MT Holder, R Vos, PE Midford, T Liebowitz, L Chan, P Hoover, T Warnow. 2009. The CIPRES Portals.

- Nixon, KC. 1999-2002. Winclada. Ithaca.
- Noble, GP, MB Rogers, PJ Keeling. 2007. Complex distribution of EFL and EF-1 alpha proteins in the green algal lineage. *Bmc Evolutionary Biology* 7.
- Posada, D, TR Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53:793-808.
- Posada, D, KA Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817-818.
- Ribichich, KF, RC Georg, SL Gomes. 2006. Comparative EST analysis provides insights into the basal aquatic fungus *Blastocladiella emersonii*. *Bmc Genomics* 7.
- Ronquist, F, JP Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
- Sakaguchi, M, K Takishita, T Matsumoto, T Hashimoto, Y Inagaki. 2009. Tracing back EFL gene evolution in the cryptomonads-haptophytes assemblage: Separate origins of EFL genes in haptophytes, photosynthetic cryptomonads, and goniomonads. *Gene* 441:126-131.
- Shimodaira, H, M Hasegawa. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17:1246-1247.
- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.
- Stamatakis, A, P Hoover, J Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Systematic Biology* 57:758-771.
- Swofford, DL. 1993. PAUP - A computer program for phylogenetic inference using Maximum Parsimony. *Journal of General Physiology* 102:A9-A9.
- Talavera, G, J Castresana. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *56:564-577*.
- Tanabe, Y, M Saikawa, MM Watanabe, J Sugiyama. 2004. Molecular phylogeny of Zygomycota based on EF-1 alpha and RPB1 sequences: limitations and utility of alternative markers to rDNA. *Molecular Phylogenetics and Evolution* 30:438-449.
- Triemer, RE, E Linton, W Shin, A Nudelman, A Monfils, M Bennett, S Brosnan. 2006. Phylogeny of the euglenales based upon combined SSU and LSU rDNA sequence comparisons and description of *Discoplastis* gen. nov (Euglenophyta). *Journal of Phycology* 42:731-740.

- Watanabe, MM, M Hiroki. 1997. NIES - Collection List of Strains. National Institute for Environmental Studies, Tsukuba 5th Ed.
- Yoon, HS, J Grant, YI Tekle, M Wu, BC Chaon, JC Cole, JM Logsdon, DJ Patterson, D Bhattacharya, LA Katz. 2008. Broadly sampled multigene trees of eukaryotes. *Bmc Evolutionary Biology* 8.
- Yutin, N, KS Makarova, SL Mekhedov, YI Wolf, EV Koonin. 2008. The deep archaeal roots of eukaryotes. *Molecular Biology and Evolution* 25:1619-1630.
- Zwickl, DJ. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Austin: The University of Texas at Austin.

Table 1.1. Organism information for sequences used in this study. A check mark in gene column denotes whether the sequence is EFL or EF1 α ; bold in GenBank/TBestDB/JGI/TriTryp database accession # columns denotes new sequences obtained in this study; dash in strain/ecotype/voucher ID column represents that that information was not available from the database. Supergroup and “group” information was taken from Adl et al. (2005) with some updates from recent literature cited.

Supergroup	“Group(s)”	Gene		Species	Strain/ Ecotype/ Voucher ID	Nucleotide Database Accession #	Protein Database Accession #
		EF1 α	EFL				
Archaea	Crenarchaeota	√		<i>Desulfurococcus mobilis</i>	DSM 2161	X73582; gDNA	CAA51984
	Crenarchaeota	√		<i>Sulfolobus solfataricus</i>	P2	apply for; AE006641; gDNA	AAK40559
	Euryarchaeota	√		<i>Halobacterium</i> sp.	NRC-1	apply for; NC_002607; gDNA	NP_281202
	Euryarchaeota	√		<i>Methanocaldococcus jannaschii</i>	DSM 2661	apply for; L77117; gDNA	AAB98308
	Euryarchaeota	√		<i>Thermoplasma acidophilum</i>	DSM 1728	apply for; AL445064; gDNA	CAC11586
Amoebozoa	Acanthamoebidae	√		<i>Acanthamoeba culbertsoni</i>	-	AY582829; gDNA	AAU94656
	Eumycetozoa/Myxogastria	√		<i>Physarum polycephalum</i>	-	AF016243; gDNA	AAB69706
	Eumycetozoa/Protostelia	√		<i>Planoprotostelium aurantium</i>	-	AF016240; gDNA	AAB69703
	Flabellinea/Dactylopodida	√		<i>Neoparamoeba branchipila</i>	-	FJ807261; gDNA	ACO50132
Archaeplastida	Chloroplastida/Charophyta/Streptophytina/Charales	√		<i>Chara australis</i>	FWAC 7144	EF551322; mRNA	ABQ81936
	Chloroplastida/Charophyta/Streptophytina/Plantae	√		<i>Arabidopsis thaliana</i>	Columbia	NM_001125992; mRNA	NP_001119464
	Chloroplastida/Charophyta/Streptophytina/Plantae	√		<i>Oryza sativa</i>	-	D63581; mRNA	BAA23658
	Chloroplastida/Charophyta/Streptophytina/Plantae (contaminated with other unknown green algae – not in rice genome)		√	<i>Oryza sativa</i>	Nipponbare	AK110624; mRNA	conceptual translation of AK110624
	Chloroplastida/Charophyta/Streptophytina/Plantae	√		<i>Physcomitrella patens</i> subsp. <i>patens</i>	Gransden 2004	XM_001763149; mRNA	XP_001763201
	Chloroplastida/Chlorophyta/Chlorodendrales		√	<i>Tetraselmis tetrathele</i>	NEPCC 500	EF551330; mRNA	ABQ81944
	Chloroplastida/Chlorophyta/Chlorophyceae		√	<i>Chlamydomonas reinhardtii</i>	CC-503 cw92 mt+	XM_001702295; mRNA	XP_001702347
	Chloroplastida/Chlorophyta/Chlorophyceae		√	<i>Chlorococcum</i> sp.	NEPCC 478	EF551323; mRNA	ABQ81937
	Chloroplastida/Chlorophyta/Chlorophyceae		√	<i>Gonium pectorale</i>	Kaneko3	AB353888; mRNA	BAF99250
	Chloroplastida/Chlorophyta/Chlorophyceae		√	<i>Scenedesmus obliquus</i>	-	apply for; mRNA SOL00000077	apply for

	Chloroplastida/Chlorophyta/Prasinophytæ		√	<i>Mesostigma viride</i>	CCMP 20469	DQ394295; mRNA	ABD58901
	Chloroplastida/Chlorophyta/Prasinophytæ		√	<i>Micromonas pusilla</i>	-	EF551326, EF551325; mRNA	conceptual translation of nucleotide consensus
	Chloroplastida/Chlorophyta/Prasinophytæ		√	<i>Nephroselmis olivacea</i>	40.89	FJ715712; mRNA	ACN59934
	Chloroplastida/Chlorophyta/Prasinophytæ		√	<i>Ostreococcus tauri</i>	OTTH0595	apply for; gii118721429; gDNA	CAL56419
	Chloroplastida/Chlorophyta/Trebouxiophyceae		√	<i>Helicosporidium</i> sp. ex <i>Simulium jonesii</i>	AT-2000	AY729488; mRNA	AAV34148
	Chloroplastida/Chlorophyta/Trebouxiophyceae		√	<i>Parachlorella kessleri</i>	211-11g	FJ715710; mRNA	ACN59932
	Chloroplastida/Chlorophyta/Ulvophyceae	√		<i>Acetabularia acetabulum</i>	-	EF551321; mRNA	ABQ81935
	Chloroplastida/Chlorophyta/Ulvophyceae	√		<i>Chaetomorpha coliformis</i>	CHR585485	FJ539143; mRNA	ACL97368
	Chloroplastida/Chlorophyta/Ulvophyceae		√	<i>Ochlochaete hystrix</i>	CCMP 2319	FJ539139; mRNA	ACL97364
	Chloroplastida/Chlorophyta/Ulvophyceae		√	<i>Ulva intestinalis</i> (syn. <i>Enteromorpha intestinalis</i>)	-	EF551324; mRNA	ABQ81938
	Chloroplastida/Chlorophyta/Ulvophyceae		√	<i>Urospora</i> sp.	CCMP 1082	FJ539140; mRNA	ACL97365
	Rhodophyceae/Bangiophyceae	√		<i>Cyanidioschyzon merolae</i>	10D	AB095182; gDNA	BAC67667
	Rhodophyceae/Bangiophyceae	√		<i>Porphyra yezoensis</i>	TU-1	AB048204; mRNA	BAB96818
	Rhodophyceae/Florideophyceae		√	<i>Chondrus crispus</i>	-	CO652990, CO652099, CO652788; mRNA	c.t. of nucleotide sequences
	Rhodophyceae/Florideophyceae		√	<i>Gracilaria changii</i>	-	DV963090, DV963156, DV964877; mRNA	c.t. of nucleotide sequences
Chromalveolata	Alveolata/Alveolates	√		<i>Cryptosporidium parvum</i>	Iowa II	XM_001388307; mRNA	XP_001388344
	Alveolata/Alveolates	√		<i>Plasmodium falciparum</i>	3D7	XM_001350246; mRNA	XP_001350282
	Alveolata/Ciliophora	√		<i>Stentor coerules</i>	-	AF056107; gDNA	AAD03262
	Alveolata/Ciliophora	√		<i>Tetrahymena pyriformis</i>	W	D11083; mRNA	BAA01856
	Alveolata/Dinoflagellata/Dinophyceae/Gymnodinophycidae		√	<i>Amphidinium carterae</i>	CCMP 1314	EU742870, mRNA	ACF28672
	Alveolata/Dinoflagellata/Dinophyceae/Gymnodinophycidae		√	<i>Karlodinium micrum</i>	CCMP 415	DQ666284; mRNA	ABG56231
	Alveolata/Dinoflagellata/Dinophyceae/Peridiniophycidae		√	<i>Heterocapsa triquetra</i>	-	AY729485; mRNA	AAV34145

	Alveolata/Dinzoa/ <i>Oxyrrhis</i>		√	<i>Oxyrrhis marina</i>	CCMP 1788	DQ659243; mRNA	ABF82158
	Alveolata/Dinzoa/Perkinsidae		√	<i>Perkinsus marinus</i>	ATCC 50983	apply for; gi 239901484; gDNA	EER18713
	Cryptophyceae		√	<i>Cryptomonas ovata</i>	-	AB332030; mRNA	BAG54795
	Cryptophyceae		√	<i>Guillardia theta</i>	-	AM183813; mRNA	CAJ73805
	Cryptophyceae		√	<i>Rhodomonas salina</i>	CCMP 1319	DQ659244; mRNA	ABF82159
	Haptophyta/Pavlovophyceae		√	<i>Pavlova lutheri</i>	-	AY729487; mRNA	AAV34147
	Haptophyta/Prymnesiophyceae/Isochrysidales		√	<i>Emiliana huxleyi</i>	CCMP 1516	apply for; JGI scaffold 14:616735-618748; gDNA	apply for
	Haptophyta/Prymnesiophyceae/Isochrysidales		√	<i>Isochrysis galbana</i>	-	AY729486; mRNA	AAV34146
	Haptophyta/Prymnesiophyceae/Prymnesiales		√	<i>Chrysochromulina</i> sp.	NIES-1333	AB377267; mRNA	BAG55224
	Stramenopiles/Bacillariophyta/Bacillariophytina		√	<i>Achnanthes kwaitensis</i>	-	AB368767; mRNA	BAG30802
	Stramenopiles/Bacillariophyta/Bacillariophytina		√	<i>Asterionella glacialis</i>	-	AB368768; mRNA	BAG30803
	Stramenopiles/Bacillariophyta/Bacillariophytina	√		<i>Phaeodactylum tricornutum</i>	CCAP 1055/1	XM_002181557; mRNA	XP_002181593
	Stramenopiles/Bacillariophyta/Coscinodiscophytina		√	<i>Detonula confervacea</i>	-	AB368771; mRNA	BAG30806
	Stramenopiles/Bacillariophyta/Coscinodiscophytina	√		<i>Thalassiosira pseudonana</i>	CCMP 1335	XM_002297414; mRNA	XP_002297450
	Stramenopiles/Bacillariophyta/Coscinodiscophytina		√	<i>Thalassiosira pseudonana</i>	CCMP 1335	XM_002292776; mRNA	XP_002292812
	Stramenopiles/Peronosporomycetes	√		<i>Phytophthora parasitica</i>	isolate 149	DQ138080; mRNA	ABA00716
	Stramenopiles/Synurales	√		<i>Mallomonas rasilis</i>	-	AY729873; gDNA	AAX09601
Excavata	Euglenozoa/Diplonemea/ <i>Diplonema</i>		√	<i>Diplonema ambulator</i>	ATCC 50223	FJ807247; mRNA	ACO50118
	Euglenozoa/Diplonemea/ <i>Diplonema</i>		√	<i>Diplonema papillatum</i>	ATCC 50162	FJ807248; mRNA	ACO50119
	Euglenozoa/Diplonemea		√	<i>Rhynchopus euleeides</i>	ATCC 50226	FJ807249; mRNA	ACO50120
	Euglenozoa/Euglenida/Euglenia	√		<i>Colacium elongatum</i>	M11	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenia	√		<i>Colacium mucronatum</i>	UTEX LB 2524	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenia	√		<i>Colacium</i> sp.	NJ	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenia	√		<i>Colacium vesiculosum</i>	UW Lazienki	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenia	√		<i>Cryptoglana pigra</i>	CCAP 1212/1	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenia	√		<i>Cryptoglana skujai</i>	SAG 10.88	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenia	√		<i>Discoplastis adunca</i>	ASW 08039	apply for; mRNA	apply for

Euglenozoa/Euglenida/Euglenea	√		<i>Euglena clara</i>	SAG 25.98	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena deses</i>	SAG 1224-23	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena exilis</i>	SAG 1224-11a	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena geniculata</i>	SAG 1224-21	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena gracilis</i>	Indiana University Z strain	X16890; mRNA	CAA34769
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena gracilis</i>	Z strain	FJ807238; mRNA	ACO50110
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena gracilis</i>	Z strain	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena granulata</i>	OR5	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena hiemalis</i>	CCAP 1224/35	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena (aka Astasia) longa</i>	-	FJ807237; mRNA	ACO50109
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena stellata</i>	IAM E-11	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Euglena viridis</i>	NJ001	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Eutreptia pertyi</i>	UTEX LB 1290	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Eutreptiella pomquetensis</i>	CCMP 1491	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Lepocinclis acus</i> var. <i>major</i>	UTEX LB 1316	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Lepocinclis ovum</i>	SAG 1244-8	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Lepocinclis spirogyroides</i>	SAG 1224-10b; TOPO clone 1 & 4	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Lepocinclis spirogyroides</i>	SAG 1224-10b; TOPO clones 38	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Lepocinclis steinii</i>	UTEX 523	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Lepocinclis tripteris</i>	UTEX 1311	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Monomorphina aenigmatica</i>	UTEX 1284	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Monomorphina pyrum</i>	UTEX 2354	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Monomorphina</i> sp.	MI 73	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Phacus orbicularis</i>	ACOI 996	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Phacus similis</i>	SAG 58.81	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Strombomonas acuminata</i>	SAG 1280-1	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Strombomonas borysteniensis</i>	NJ Stromb 10	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Strombomonas napiformis</i>	NJ Strombo 2	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Strombomonas ovalis</i>	NJ S115	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Strombomonas triquetra</i>	NJ S116	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Trachelomonas conradi</i>	UTEX 1313	apply for; mRNA	apply for
Euglenozoa/Euglenida/Euglenea	√		<i>Trachelomonas</i>	SAG 1283-22	apply for; mRNA	apply for

				<i>echinata</i>			
	Euglenozoa/Euglenida/Euglenea	√		<i>Trachelomonas lefevrei</i>	SAG 1283-10	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenea	√		<i>Trachelomonas planctonica</i>	Woongghi Trach 1 (WT1)	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenea	√		<i>Trachelomonas similis</i>	SAG 1283-14	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Euglenea	√		<i>Trachelomonas volvocina</i>	UTEX 1327	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Heteronematina	√		<i>Entosiphon</i> sp.	Carolina Biol.	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Heteronematina	√		<i>Entosiphon sulcatum</i>	-	FJ807254; gDNA	ACO50125
	Euglenozoa/Euglenida/Heteronematina	√		<i>Peranema trichophorum</i>	CCAP 1260/1b	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Heteronematina	√		<i>Peranema trichophorum</i>	CCAP 1260/1b	FJ807243; mRNA	ACO50115
	Euglenozoa/Euglenida/Heteronematina	√		<i>Ploeotia costatum</i>	CCAP 1265/1	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Heteronematina	√		<i>Ploeotia oblonga</i>	ATCC 50573	apply for; mRNA	apply for
	Euglenozoa/Euglenida/Heteronematina	√		<i>Ploeotia oblonga</i>	ATCC 50573	apply for; gDNA	apply for
	Euglenozoa/Kinetoplastida/Metakinoplastina/Eubodonida	√		<i>Bodo saltans</i>	CCAP 1907/2	FJ807252; gDNA	ACO50123
	Euglenozoa/Kinetoplastida/Metakinoplastina/Neobodonida	√		<i>Dimastigella trypaniformis</i>	ATCC 50263	FJ807253; gDNA	ACO50124
	Euglenozoa/Kinetoplastida/Metakinoplastina/Neobodonida		√	<i>Neobodo saliens</i> (syn. <i>Bodo saliens</i>)	ATCC 50358	FJ807267; mRNA	ACO50138
	Euglenozoa/Kinetoplastida/Metakinoplastina/Neobodonida	√		<i>Rhynchobodo</i> sp.	ATCC 50359	FJ807264; gDNA	ACO50135
	Euglenozoa/Kinetoplastida/Metakinoplastina/Neobodonida	√		<i>Rhynchomonas nasuta</i>	ATCC 50292	FJ807253; mRNA	ACO50121
	Euglenozoa/Kinetoplastida/Metakinoplastina/Parabodonida		√	<i>Trypanoplasma borreli</i>	Tt-JH	FJ807268; mRNA	ACO50139
	Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	√		<i>Blastocrithidia culicis</i>	ATCC 30268	FJ807251; gDNA	ACO50122
	Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	√		<i>Herpetomonas muscarum</i>	ATCC 30260	FJ807255; gDNA	ACO50126
	Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	√		<i>Herpetomonas pessoai</i>	ATCC 30252	FJ807256; gDNA	ACO50127
	Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	√		<i>Leishmania braziliensis</i>	MHOM/BR/ 75/M2904	XM_001563727; mRNA	XP_001563777
	Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	√		<i>Leishmania infantum</i>	JPCM5	XM_001464627; mRNA	XP_001464664
	Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	√		<i>Leishmania major</i>	Friedlin	XM_001682206; mRNA	XP_001682258
	Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	√		<i>Leishmania tarentolae</i>	Parrot-II	apply for; gDNA TriTrypDB r1596.contig933	apply for
	Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	√		<i>Leishmania tarentolae</i>	UC	FJ807257; gDNA	ACO50128
	Euglenozoa/Kinetoplastida/Metakinoplastina/	√		<i>Leptomonas bifurcata</i>	-	FJ807258; gDNA	ACO50129

	Trypanosomatida						
	Euglenozoa/Kinetoplastida/Metakinoplastina/ Trypanosomatida	√		<i>Leptomonas costaricensis</i>	-	FJ807259; gDNA	ACO50130
	Euglenozoa/Kinetoplastida/Metakinoplastina/ Trypanosomatida	√		<i>Leptomonas podlipaevi</i>	-	FJ807260; gDNA	ACO50131
	Euglenozoa/Kinetoplastida/Metakinoplastina/ Trypanosomatida	√		<i>Trypanosoma avium</i>	-	FJ807265; gDNA	ACO50136
	Euglenozoa/Kinetoplastida/Metakinoplastina/ Trypanosomatida	√		<i>Trypanosoma brucei</i>	TREU927	XM_817372; mRNA	XP_822465
	Euglenozoa/Kinetoplastida/Metakinoplastina/ Trypanosomatida	√		<i>Trypanosoma brucei equiperdum</i>	STIB818	FJ807266; gDNA	ACO50137
	Euglenozoa/Kinetoplastida/Metakinoplastina/ Trypanosomatida	√		<i>Trypanosoma cruzi</i>	CL Brener	XM_814346; mRNA	XP_819439
	Euglenozoa/Kinetoplastida/Metakinoplastina/ Trypanosomatida	√		<i>Trypanosoma rangeli</i>	San Agustin clone 1	AY785727; gDNA	AAX40412
	Euglenozoa/Kinetoplastida/Prokinoplastina	√		<i>Perkinsiella amoebae</i>	from host <i>Neoparamoeba branchiphila</i> (AMOP1)	FJ807262; gDNA	ACO50133
	Euglenozoa/ <i>Incertae sedis</i> (Ma 2005; Yubuki et al. 2009)		√	<i>Petalomonas cantuscygni</i>	CCAP 1259/1	FJ807263; mRNA	ACO50134
	Fornicata/Eopharyngia/Diplomonadida	√		<i>Hexamita inflata</i>	AZ-4	U37081, U93087; gDNA	c.t. of nucleotide sequences
	Fornicata/Eopharyngia/Diplomonadida	√		<i>Spironucleus vortens</i>	ATCC 50386	U94406; gDNA	AAB81020
	Heterolobosea/Acrasidae	√		<i>Acrasis rosea</i>	T-235	AF190771; gDNA	AAG48934
	Heterolobosea/Gruberellidae	√		<i>Stachyamoeba lipophora</i>	-	FJ807246; mRNA	ACO50140
	Heterolobosea/Vahlkampfiidae	√		<i>Naegleria andersoni</i>	A2	AF058283; gDNA	AAD15802
	Heterolobosea/Vahlkampfiidae	√		<i>Naegleria gruberi</i>	NEG-M	DQ295229; gDNA	ABC54650
	Heterolobosea/Vahlkampfiidae	√		<i>Sawyeria marylandensis</i>	ATCC 50653	apply for; mRNA SML00000003	apply for
	Jakobida/Andalucia	√		<i>Andalucia incarcerationata</i>	MB1, Canada	EU334884; gDNA	ACA34537
	Jakobida/Histionidae	√		<i>Histiona aroides</i>	ATCC 50634	FJ807241; mRNA	ACO50111
	Jakobida/Histionidae	√		<i>Reclinomonas americana</i>	ATCC 50283	FJ807239; mRNA	ACO50135
	Jakobida/Jakoba	√		<i>Jakoba libera</i>	ATCC 50422	FJ807244; mRNA	ACO50113
	Jakobida/ <i>Incertae sedis</i>	√		<i>Seculamonas ecuadoriensis</i>	ATCC 50688	FJ807245; mRNA	ACO50117
	<i>Malawimonas</i>	√		<i>Malawimonas californiana</i>	ATCC 50740	FJ807242; mRNA	ACO50114
	<i>Malawimonas</i>	√		<i>Malawimonas jakobiformis</i>	ATCC 50310	DQ295227; gDNA	ABC54649
	Parabasalia	√		<i>Histomonas meleagridis</i>	-	FM200063; mRNA	CAQ86674
	Parabasalia	√		<i>Trichomonas vaginalis</i>	clone G3	XM_001328923; mRNA	XP_001328958
	Preaxostyla/Oxymonadida	√		<i>Dinenympha exilis</i>	A-5	AB007028; mRNA	BAA22607

	Preaxostyla/Oxymonadida	√		<i>Monocercomonoides</i> sp.	clone PA	EF474125; mRNA	ABU93326
	Preaxostyla/Oxymonadida	√		<i>Streblomastix strix</i>	clone 1	AY188861; gDNA	AAO46118
	Preaxostyla/ <i>Trimastix</i>	√		<i>Trimastix pyriformis</i>	ATCC 50562	DQ295235; gDNA	ABC54653
Opisthokonta	Choanomonada		√	<i>Monosiga brevicollis</i>	ATCC 50154	AY026073; mRNA	AAK27413
	Choanomonada	√		<i>Monosiga ovata</i>	-	AY582824; mRNA	AAU94651
	Choanomonada	√		<i>Salpingoeca amphoridium</i>	-	DQ059027; gDNA	AAU99757
	Fungi/Ascomycota	√		<i>Neurospora crassa</i>	OR74A	XM_959775; mRNA	XP_964868
	Fungi/Ascomycota	√		<i>Saccharomyces cerevisiae</i>	-	X01638; gDNA	CAA25798
	Fungi/Basidiomycota	√		<i>Cryptococcus neoformans</i>	JEC21	XM_568462; mRNA	XP_568462
	Fungi/Basidiomycota	√		<i>Schizophyllum commune</i>	1-40	X94913; gDNA	CAA64399
	Fungi/Chytridiomycetes/Blastocladales/Blastocladiaceae		√	<i>Allomyces macrogynus</i>	-	apply for; mRNA AML00000823	apply for
	Fungi/Chytridiomycetes/Blastocladales/Blastocladiaceae		√	<i>Blastocladiella emersonii</i>	-	EF064246; mRNA	ABK34893
	Fungi/Chytridiomycetes/Blastocladales/Catenariaceae		√	<i>Catenophylctis</i> sp.	JEL298	DQ275345; gDNA	ABB84540
	Fungi/Chytridiomycetes/Blastocladales/Coelomomycetaceae		√	<i>Coelomomyces stegomyiae</i>	AFTOL-ID 18	DQ275335; gDNA	ABB84530
	Fungi/Chytridiomycetes/Monoblepharidales	√		<i>Harpochytrium</i> sp.	JEL94	AF450113; gDNA	AAL87075
	Fungi/Chytridiomycetes/Monoblepharidales	√		<i>Monoblepharis insignis</i>	S9-7	AF450112; gDNA	AAL87074
	Fungi/Chytridiomycetes/Spizellomycetales		√	<i>Gaertneriomyces semiglobiferous</i>	AFTOL-ID 34	DQ275338; gDNA	ABB84533
	Fungi/Chytridiomycetes/Spizellomycetales	√		<i>Rhizophlyctis harderi</i>	JEL171	AF450114; gDNA	AAL87076
	Fungi/Chytridiomycetes/Spizellomycetales		√	<i>Spizellomyces punctatus</i>	-	apply for; mRNA SPL00000880	apply for
	Fungi/Zygomycota/ <i>Basidiobolus</i>		√	<i>Basidiobolus ranarum</i>	NRRL34594; AFTOL-ID 301	DQ275340; gDNA	ABB84535
	Fungi/Zygomycota/Dimargaritales	√		<i>Absidia glauca</i>	CBS 101.48	X54730; gDNA	CAA38529
	Fungi/Zygomycota/Entomophthora		√	<i>Conidiobolus coronatus</i>	AFTOL-ID 137	DQ275337; gDNA	ABB84532
	Fungi/Zygomycota/Entomophthora		√	<i>Entomophthora muscae</i>	AFTOL-ID 28	DQ275343; gDNA	ABB84538
	Fungi/Zygomycota/Harpellales	√		<i>Smittium simulii</i>	CAL81	AY582822; gDNA	AAU94649
	Mesomycetozoa	√		<i>Amoebidium parasiticum</i>	-	AY582828; gDNA	AAU94655
	Mesomycetozoa	√		<i>Capsaspora owczaraki</i>	-	DQ403163; mRNA	ABD67497
	Mesomycetozoa		√	<i>Creolimax fragrantissima</i>	Pe4 from host <i>Phascolosoma agassizii</i>	EU169924- EU169930; gDNA	c.t. of nucleotide sequences

	Mesomycetozoa		√	<i>Sphaeroforma arctica</i>	-	DQ403164; mRNA	ABD67498
	Metazoa/Animalia	√		<i>Caenorhabditis elegans</i>	Bristol N2	NM_066119; mRNA	NP_498520
	Metazoa/Animalia	√		<i>Drosophila melanogaster</i>	-	NM_165850; mRNA	NP_725085
	Metazoa/Animalia	√		<i>Ornithorhynchus anatinus</i>	Glennie (liver, lung, brain, spleen), ♀	XM_001507841; mRNA	XP_001507891
	Metazoa/Animalia	√		<i>Xenopus laevis</i>	oocyte	NM_001086133; mRNA	NP_001079602
	Metazoa/Animalia		√	<i>Xenopus laevis</i>	Intestine, metamorphosis - Niewkoop and Faber stages 56 to 62 (tadpoles), pooled, normal, undetermined sex	NM_001097658; mRNA	NP_001091127
Rhizaria	Chlorarachniophyta		√	<i>Chlorarachnion reptans</i>	NEPCC 449 (cytosolic B)	EU810315; mRNA	ACF24583
	Chlorarachniophyta		√	<i>Gymnochlora stellata</i>	CCMP 2057 (cytosolic A)	EU810316; mRNA	ACF24584
	Chlorarachniophyta		√	<i>Lotharella amoebiformis</i>	CCMP 2058	EU810318; mRNA	ACF24586
	Foraminifera		√	<i>Planoglabratella operculis</i>	-	AB334123; mRNA	BAG54798
	Foraminifera		√	<i>Reticulomyxa filosa</i>	-	EU810334; mRNA	ACF24599
Others	periplastid compartment-targeted Chlorarachniophyta		√	in <i>C. reptans</i>	NEPCC 449	EU810327; mRNA	ACF24594
	periplastid compartment-targeted Chlorarachniophyta		√	in <i>G. stellata</i>	CCMP 2057	EU810329; mRNA	ACF24595
	periplastid compartment-targeted Chlorarachniophyta		√	in <i>L. amoebiformis</i>	CCMP 2058	EU810330; mRNA	ACF24596
	uncultured eukaryote from environmental sample		√	environmental sample	chlorarachniophyte cell pellet	EU810323; mRNA	ACF24591

Table 1.2. The detailed EF1 α -extended (annealing temperature (AT) was set to either 47 or 52°C) and standard EF1 α PCR programs (AT was set to either 37 or 50.4°C).

Step Number	EF1 α -extended PCR program Step Details	EF1 α PCR programs Step Details
1	94.0 for 2:00	95.0 for 5:00
2	2.0/s to 94.0	95.0 for 0:30
3	94.0 for 0:20	AT for 0:30
4	2.0/s to 70.0	72.0 for 2:00
5	70.0 for 0:01	Goto 2, 29 times
6	0.1/s to AT	72.0 for 10:00
7	AT for 0:30	10.0 for ever
8	0.2/s to 72.0	END
9	72.0 for 1:30	
10	Goto 2, 9 times	
11	2.0/s to 94.0	
12	94.0 for 0:20	
13	2.0/s to 70.0	
14	70.0 for 0:01	
15	0.1/s to AT	
16	AT for 0:30	
17	0.2/s to 72.0	
18	72.0 for 1:30 + 5sec/cycle	
19	Goto 11, 9 times	
20	2.0/s to 94.0	
21	94.0 for 0:20	
22	2.0/s to 70.0	
23	70.0 for 0:01	
24	0.1/s to AT	
25	AT for 0:30	
26	0.2/s to 72.0	
27	72.0 for 2:15 + 20sec/cycle	
28	Goto 20, 14 times	
29	72.0 for 3:00	
30	10.0 for ever	
31	END	

Table 1.3. Degenerate primer names, origins and sequences used in this study. A lowercase “i” denotes an inosine. Bold glycine (G) and cytosine (C) residues indicate that they were added to the end of the primer sequence to increase efficiency of the PCR products into the cloning vectors. Underlined text denotes hairpin sequence S. Baldauf introduced into the primer sequence that is not specific to the EF1 α gene.

Primer Name	Original Primer Author	Original Primer Sequence (5' to 3')	Sequence for this study (5' to 3')	Corresponding Protein Motif
EUG EF1a 1F	S. Jardeleza	this study	G GGi AAR GAi AAR GTi CAY ATN ARY YT	GK(D/E)KVH(I/M)(S/N)
1F_FL	S. Jardeleza	this study	G GGV AAR GAi AAR GTG CAY ATi ARY YT	GK(D/E)KVH(I/M)(S/N)
EUG EF1a 1R	S. Jardeleza	this study	NCC NAR iGG iGS RTA RTC iKT RAA	F(T/N/?)DYPPLG
EUG EF1a 1Rgg	S. Jardeleza	this study	GG NCC NAR iGG iGS RTA RTC iKT RAA	F(T/N/?)DYPPLG
EUG EF1a 2R	S. Jardeleza	this study	CC NAC NGC iAC iTG YYG iCG CAT RTC	DMRQTVAVG
GG_EF1a_F3	G. Gile	TTC GAG AAG GAG GCN GCN GAR ATG AA	GG TTY GAG AAG GAR GCi DVY GAi ATS GG	FEKEAAEMN
Bal_7R2	S. Baldauf	GAT GAA ACC AAC GTT RTC NCC NGG	G GTT GAA RCC iAC RTT GTC VCC NGG	PGDNVGFN
Bal_10R	S. Baldauf	<u>CAG</u> AAT TCC GAC AGC NAC NTG YTG	C <u>CAG</u> AAT TCC GAC AGC NAC NGTYTG	QTVAVGI

Table 1.4. Summary of alignments, which all contained 194 operational taxonomic units (OTUs). Alignment short names are abbreviations of the alignment process and type of positions included in the alignment as detailed in separate columns of the table, shown with an “X” in the appropriate columns. Mo=aligned with MUSCLE and edited by eye against amino acid and previously published alignments of EF1 α and/or EFL; lc=phylogenetically informative indels excised from the original alignment by GBlocks were replaced; Go=original alignment was run through GBlocks and any unsupported positions were excluded in the resulting alignment; 3rd=alignments with the 3rd codon nucleotide positions removed in the EF1 α sequences. Nucleotide model of sequence evolution (MoSE) found to be preferred by both hLRT and AIC in Modeltest for each alignment is given, whereas amino acid MoSE was determined by AIC in ProtTest. In amino acid MoSE, “F” denotes that the frequency of amino acid usage was estimated directly from the alignment. Number of base pairs (bp) in each nucleotide alignment or amino acid (aa) residues and the number of parsimony informative characters (Pars. Inform) detected in Winclada for each alignment are also described, except for amino acid alignments because MP was not performed on them.

Short Name	Mo	lc	Go	3 rd	MoSE	bp/aa	Pars. Inform.
Mo	X				GTR+I+ Γ	1803	1261
MolcGo	X	X			GTR+I+ Γ	1389	1163
MoGo	X		X		GTR+I+ Γ	1248	1042
Mo3rd	X			X	GTR+I+ Γ	1202	771
MolcGo3rd	X	X		X	GTR+I+ Γ	926	706
MoGo3rd	X		X	X	GTR+I+ Γ	832	628
aaMo	X				RTREV+I+ Γ +F	601	no MP
aaMolcGo	X	X			RTREV+I+ Γ +F	451	no MP
aaMoGo	X		X		RTREV+I+ Γ +F	385	no MP

Table 1.5. Summary of MP, ML and MLE nucleotide analyses. Single ML tree (best $-\ln L$ score) for the ML and MLE analyses and any MP trees with the best parsimony score were used in the topology test with the MolcGo3rd alignment. AU rank is from CONSEL (*=tip-level polytomies in original best tree forced to bifurcate based on previous studies to use in the AU test). Trees ranked 19-57 were statistically rejected (p value <0.05). The number of trees (# trees) used to create the support tree for each analysis is stated.

Analysis & Parameters	Alignment Input	AU rank					# trees
<u>TNT (MP)</u> New Technology Search: Sectorial Search, Ratchet, Drift & Tree fusing transversion cost 2:1 standard bootstrapping	Mo	21	46	48	47	1000	
	MolcGo	52			56		1000
	MoGo	44				1000	
	Mo3rd	25	22	26	23	37	1000
	MolcGo3rd	27	33	24	42	34	1000
	MoGo3rd	36	45	39	30	32	1000
<u>RAxML (ML)</u> autodetect bootstrap replicates GTR+I+Γ	Mo	20				250	
	MolcGo	28				300	
	MoGo	50				250	
	Mo3rd	3				300	
	MolcGo3rd	1				300	
	MoGo3rd	8				300	
<u>GARLI (ML)</u> defaults on website for GTR+I+Γ, with 6cats, where invariant sites & statefreq were estimated empirically from alignment	Mo	40				100	
	MolcGo	31				100	
	MoGo	57				100	
	Mo3rd*	11				100	
	MolcGo3rd*	5				100	
	MoGo3rd*	16				100	
<u>MrBayes (MLE)</u> McMC parameters: nchains=4; Temp=0.1; Swapfreq=1; Nswaps=1 GTR+I+Γ; 4 gamma categories; est. empirical freq. 30 mill gens with 15 mill burn-in; 4 runs each with trees after burn-in combined to make one .con tree	Mo	53	38	35	51	60,000	
	MolcGo	41	19	29	43	60,000	
	MoGo	54	49	55	58	60,000	
	Mo3rd	13	9	14	18	60,000	
	MolcGo3rd	4	6	7	3	60,000	
	MoGo3rd	17	10	15	12	60,000	

Table 1.6. Summary of ML and MLE amino acid phylogenetic analyses. The number of trees (# trees) used to create the support tree for each analysis is stated.

Analysis & Parameters	Alignment Input	# trees
<u>RAxML (ML)</u> autodetect bootstrap replicates RTREV+I+Γ+F	aaMo	350
	aaMolcGo	300
	aaMoGo	350
<u>MrBayes (MLE)</u> McMC parameters: nchains=4; Temp=0.2; Swapfreq=1; Nswaps=1 RTREV+I+Γ+F; 4 gamma categories; est. empirical freq. 6 million gens with 3 million burn-in; 6 runs each with trees after burn-in combined to make one .con tree	aaMo	180,000
	aaMolcGo	180,000
	aaMoGo	180,000

Table 1.7. Summary of recovered clades from all phylogenetic analyses. Analyses types were TNT (“T”; MP; 1000bs MP tree), RAxML (“R”; ML; “best tree” used with rapid bootstrap support), GARLI (“G”; ML; consensus tree), and MrBayes (“M”; MLE; consensus tree). The following subcategories are reported because they have been recovered as clades in previous studies: unikonts, bikonts, CMM (Choanomonada + Mesomycetozoa (specifically Ichthyosporae) + Metazoa), “Group 1” Excavates, “Group 2” Excavates, “Group 3” Excavates, SARs (Stramenopila + Alveolata + Rhizaria and perhaps includes the Cryptophyceae + Haptophyta pairing). EFL Archaeplastida is assumed to include the Chlorarachniophyta EFL copy targeted to the periplastid compartment and CMM includes the uncultured eukaryote species EFL sequence that has consistently branched with that clade in previous studies. *O. sativa* EFL is not taken into consideration as a Charophyte/Streptophytina clade-member because it is deemed an EFL sequence of some unknown green algal contaminant as the sequence is not recovered in the *O. sativa* genome sequence databases. Tallies of the number of recovered clades for each analysis type are at the bottom of each alignment type column. A * denotes the highest number of recovered clades for that data set type for each analysis type, with both values marked in cases of ties.

Clade	DNA Alignment						Protein Alignment		
	Mo	MolcGo	MoGo	Mo3rd	MolcGo3rd	MoGo3rd	Mo	MolcGo	MoGo
aEF1α	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
*Crenarchaeota	TRGM	TRGM	RGM	TRGM	TRGM	TRGM	RM	RM	RM
*Euryarchaeota	RM	RGM	RGM	TRGM	TRGM	TRGM	RM	RM	RM
eEF1α	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
* “unikonts”				RM	RM	RM			
**Amoebozoa	M	M	RM	RM	RM	RM	M	M	M
***Eumycetozoa	TRGM	TRGM	TRM	RGM	RGM	RGM	RM	RM	RM
**Opisthokonta		G		RGM	RGM	RGM	RM	RM	RM
***Fungi		TG		RM	RGM	RM	RM	RM	RM
****Ascomycota									
****Basidiomycota		TRGM	T		T	T		RM	M
****Chytridiomycetes	T	T	T	T	T	T			
*****Monoblepharidales									
****Zygomycota									
***CMM	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
****Choanomonada	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
****Mesomycetozoa/Ichthyosporae	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
****Metazoa/Animalia	TRGM	TRM	TR	T			RM	RM	
* “bikonts”									
**Archaeplastida									
***Chloroplastida	RGM	RGM	RGM	TRGM	RGM	RGM	RM	RM	RM
****Charophyta/Streptophytina	RGM	RGM	RGM	RGM	TRGM	TRGM	RM	RM	RM

****Chlorophyta/Ulvophyceae	RGM	RGM	RGM	TRGM	RGM	RGM	RM	RM	RM
***Rhodophyceae/Bangiophyceae	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
**Chromalveolata									
**Stramenopila	R	R				R	RM	RM	RM
***Bacillariophyta	RGM	RGM	TRGM	RM	RGM	RGM	RM	RM	M
**Alveolata									
***Apicomplexa	RGM	RGM	RGM	TRGM	TRGM	TRGM	RM	RM	RM
***Ciliophora						R			
**Excavata									
*** “Group 1”	RGM	TRGM	TRGM	M	RGM	M			
****Fornicata/Eopharyngia/Diplomonadida	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
****Parabasalia	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
*** “Group 2”									
**** <i>Malawimonas</i>	RGM	RGM	RGM	RGM	RGM	RGM	M	M	RM
****Preaxostyla	RGM	RGM	RGM	RGM	RGM	TRGM	RM	RM	RM
*****Oxymonadida	RGM	TRGM	TRGM	RGM	RGM	RGM	M	M	
*** “Group 3”; “Discoba”									
****Euglenozoa	R	G		RGM	RGM	RGM	RM	M	RM
*****Euglenida	R	G		RGM	RGM	RGM	RM	RM	RM
*****Euglenea									
*****Heteronematina									
****Kinetoplastida	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
*****Metakinoplastida									
*****Neobodonida									
*****Trypanosomatida	RGM	RGM	RGM	RGM	RGM	RGM	RM	M	M
***Heterolobosea									
*****Vahlkampfiidae	G	RM	RG	RGM	RGM	RG			
****Jakobida	R	RM	R	RGM	RGM	RGM	RM	M	RM
*****Histionidae	TRGM			TRGM	TRGM	TRGM	RM		RM
EFL	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
* “unikonts”									
***Opisthokonta									
****Fungi									
*****Chytridiomycetes	GM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
*****Blastocladales		RM	RGM	R	R	RGM		M	RM
*****Blastocladiaceae	T	T	TRGM	T	T	T			
*****Spizellomycetales	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
*****Zygomycota	TRGM	TRGM	TRGM	TRM	TRM	TRM			

*****Entomophthora									
****CMM	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
*****Mesomycetozoa/Icthyosporia	TRGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
*“bikonts”									
**SARs									
***Stramenopila/Bacillariophyta	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
****Bacillariophytina	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	
****Coscinodiscophytina	TRG		RGM						
***Alveolata/Dinoflagellata	RGM	RGM	RG		T	T			
****Dinoflagellata/Dinophyceae	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
*****Gymnodinophyceae	T	T	T						
***Rhizaria							RM	RM	RM
****Chlorarachniophyta	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
****Foraminifera	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
***Cryptophyceae + Haptophyta									
****Cryptophyceae	TRGM	TRGM	TRM	TRGM	TRGM	TRGM	RM	RM	RM
****Haptophyta	RGM	RGM	RGM	RGM	RGM	R	RM	RM	RM
*****Prymnesiophyceae	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
*****Isochrysidales	T	T	T	T	T	T			
***Archaeplastida									
****Chloroplastida									
*****Chlorophyta									
*****Chlorophyceae									
*****Prasinophytaceae									
*****Trebouxiophyceae									
*****Ulvophyceae									
****Rhodophyceae/Florideophyceae	TRGM	TRGM	TRGM	TRGM	TRGM	TRGM	RM	RM	RM
***Excavata/“Group 3”/Euglenozoa									
*****Diplonemea	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
*****Diplonema	RM	RM	RGM	RM	RGM	RGM			
*****Kinetoplastida/Metakinetoplastida	RGM	RGM	RGM	RGM	RGM	RGM	RM	RM	RM
Total clades recovered	22T 47R 43G 44M	23T* 47R 45G 47M	22T 48R 43G 44M	22T 49R 42G 49M*	22T 50R 46G* 49M*	23T* 51R* 44G 48M	44 R* 47M	41 R 48M*	42 R 45M

Figure 1.1. The Mo nucleotide alignment of 194 OTUs. The OTUs are grouped according first to gene, then to organisms that were related in previous literature and are given a number from 001-194 which is consistent throughout the pages of the alignment. The column positions, from 1-1803, are given in ten base increments at the top of each page. Questionably named, identified or contaminated sequences are in quotation marks. The conceptual translation based on the universal genetic code is given under each triplet for each sequence to indicate the approximate amino acid translation for each respective sequence; however, some of these will be different from the actual amino acid sequences if the given organism has a different genetic code. The MolcGo alignment excludes the following columns numbers, with blocks excluded in brackets: [1-54], [550-573], [634-636], [658-714], [805-816], [829-834], [991-993], [1108-1119], [1252-1260], [1264-1267], [1273-1275], [1288-1299], [1333-1335], [1357-1389], [1480-1488], [1504-1515], [1558-1560], [1669-1803]. The MoGo alignment excludes all columns that the MolcGo alignment does as well as the following additional columns, with blocks excluded in brackets: [433-456], [643-713], [721-738], [844-900], [931-933], [1093-1107], [1210-1212], [1453-1461], [1642-1656]. The third codon positions were removed from those three alignment types to create Mo3rd, MolcGo3rd and MoGo3rd, respectively.

	10	20	30	40	50	60	70	80	90	100	110	
001_aEF1a_Desulfurococcus_mobilis				-----ATGA	GTGCACCGCA	G---AAGCCA	CATTTAAAT	TAGTGATAAT	AGGGCACGTA	GACCACGGTA	AAAGCACGAT	[71]
002_aEF1a_Sulfolobus_solfataricus				-----	--ATGCTCTA	A---AAGCCT	CACCTTAAT	TAATAGTAAT	AGGTCATATC	GATCATGGTA	AAAGCACACT	[65]
003_aEF1a_Thermoplasma_acidophilum				-----a	TgGcAaGtCa	G---AaAcCa	CaCtTgAaTc	TgAtAaCgAt	AgGtCaTgTa	GaTcAYGgAa	AaTcTaCaTt	[68]
004_aEF1a_Halobacterium_sp.				-----	--ATGAGCGA	C---AACC GC	CACCAGAACC	TGGCCGTCAT	CGGCCACGTC	GACCACGGCA	AGAGCACGAT	[65]
005_aEF1a_Methanocaldococcus_jannaschii				-----A	TGGCAAGCA	A---AAACCA	GTATTAAACG	TAGCATTTCA	TGGACACGTA	GATGCAGGTA	AGTCAACAAC	[68]
006_EF1a_Acanthamoeba_culbertsoni				-----			-----G	AGCTCGTTAT	TGGGCACGTG	GATGCCGGTA	AGTCGACCAC	[41]
007_EF1a_Neoparamoeba_branchipila				-----						GATGCCGGAA	AGTCGACCAC	[20]
008_EF1a_Planoprotostelium_aurantium				-----						---GCTGGAA	AGTCGACCAC	[17]
009_EF1a_Physarum_polycephalum				-----						---GCTGGCA	AGTCGACCAC	[17]
010_EF1a_Amoebidium_parasiticum				-----				--CTCGTTAT	TGGGCATGTA	GA CTCCGGCA	AGTCGACCCT	[38]
011_EF1a_Capsaspora_owczaraki				-----A	TGGGCAAGGA	G---AAGACT	CACATTAACA	TCGTCGTTAT	CGGCCACGTC	GATTCCGGCA	AGTCGACCCT	[68]
012_EF1a_Monosiga_ovata				-----			-----G	GATCCGTTAT	TGGGCATGTA	GA CTCCGGCA	AGTCGACCCT	[41]
013_EF1a_Salpingoeca_amphoridium				-----			-----GTTAT	TGGACATGTG	GATTCCGGCA	AGTCGACCAC		[35]
014_EF1a_Ceanorhabditis_elegans				-----A	TGGGAAAAGA	A---AAGGTT	CATATCAACA	TCGTCGTCAT	CGGCACGTC	GACTCCGGAA	AGTCGACCAC	[68]
015_EF1a_Drosophila_melanogaster				-----A	TGGGCAAGGA	A---AAGATT	CACATTAACA	TTGTCGTGAT	CGGCACGTC	GATTCCGGTA	AGTCGACCAC	[68]
016_EF1a_Ornithorhynchus_anatinus				-----A	TGGGAAAAGA	A---AAGACC	CACATCAACA	TCGTGGTCAT	TGGACATGTT	GACTCTGGAA	AATCCACCAC	[68]
017_EF1a_Xenopus_laevis				-----A	TGGGAAAAGA	A---AAGATT	CACATTAACA	TCGTTGTGAT	CGGGCATGTG	GA CTCCGGCA	AGTCGACCAC	[68]
018_EF1a_Saccharomyces_cerevisiae				-----A	TGGGTAAGA	G---AAGTCT	CACATTAACG	TTGTCGTTAT	CGGTCATGTC	GATTCTGGTA	AGTCTACCAC	[68]
019_EF1a_Neurospora_crassa				-----A	TGGGCAAGGA	GGACAAGACT	CACATCAACG	TCGTCGTTAT	CGGCCACGTC	GATTCCGGCA	AGTCTACCAC	[71]
020_EF1a_Cryptococcus_neoformans				-----A	TGGGTAAGGA	C---AAGCTG	CACGTCACGC	TCGTTGTTAT	CGGTCACGTC	GA CTCCGGTA	AGTCGACCAC	[68]
021_EF1a_Schizophyllum_commune				-----A	TGGGAAAAGA	G---AAGCTC	CACGTTAAACG	TCGTCGTCAT	CGGTCACGTC	GA CTCCGGCA	AGTCGACCAC	[68]
022_EF1a_Monoblepharis_insignis				-----								[0]
023_EF1a_Rhizophlyctis_harderi				-----								[0]
024_EF1a_Harpochytrium_sp.				-----								[0]
025_EF1a_Absidia_glauca				-----A	TGGGTAAGA	A---AAGACT	CACGTTAAACG	TCGTTGTGAT	TGGTCACGTC	GATTCTGGTA	AATCCACCAC	[68]
026_EF1a_Smittium_simulii				-----			-----G	AGCTCGTTAT	TGGACACGTC	GATTCCGGTA	AATCTACCAC	[41]
027_EF1a_Chara_australis				-----						---TCTGGCA	AGTCGACGAC	[17]
028_EF1a_Physcomitrella_patens				-----A	TGGGTAAGGA	G---AAGGTG	CACATCAATA	TCGTGGTCAT	CGGCCATGTC	GA CTCTGGCA	AGTCGACGAC	[68]
029_EF1a_Acetabularia_acetabulum				-----						---TCTGGTA	AGTCGACCAC	[17]
030_EF1a_Chaetomorpha_coliformis				-----								[0]
031_EF1a_Arabidopsis_thaliana				-----A	TGGGTAAGA	G---AAGTTT	CACATTAACA	TTGTGGTCAT	TGGTCATGTT	GATTCTGGAA	AGTCGACCAC	[68]
032_EF1a_Oryza_sativa				-----A	TGGGTAAGGA	G---AAGACG	CACATCAACA	TTGTGGTCAT	TGGCCACGTC	GA CTCCGGCA	AGTCGACCAC	[68]
033_EF1a_Cyanidioschyzon_merolae				-----A	TGGGTAAGGA	A---AAGACG	CACGTTTCCG	TGGTGGTTAT	CGGTCACGTC	GA CTCCGGCA	AGTCGACTAC	[68]
034_EF1a_Porphyrha_yezoensis				-----A	TGGGGAAGGA	G---AAGCAG	CATGTGTCCA	TTGTGGTCAT	TGGCCACGTC	GA CTCTGGCA	AGTCGACGAC	[68]
035_EF1a_Plasmodium_falciparum				-----A	TGGGGAAGGA	A---AAAACA	CATATTAACCT	TAGTTGTTAT	TGGTCACGTC	GATGTGGGAA	AATCAACCAC	[68]
036_EF1a_Cryptosporidium_parvum				-----A	TGGGTAAGGA	A---AAGACT	CACATTAACCT	TGGTCGTTAT	TGGTCACGTC	GA CTCTGGTA	AATCAACCAC	[68]
037_EF1a_Tetrahymena_pyriformis				-----A	TGGCTAGAGG	TGATAAGGTT	CATATTAACCT	TGGTCGTTAT	CGGTCACGTC	GACTCCGGTA	AGTCGACCAC	[71]
038_EF1a_Stentor_coeruleus				-----				-----GTG	GATTCCGGTA	AGTCGACTTC		[23]
039_EF1a_Phytophthora_parasitica				-----A	TGGGCAAGGA	A---AAGGTT	CACATTTGCG	TGGTCGTGAT	TGGCCACGTC	GACGCCGGTA	AGTCGACCAC	[68]
040_EF1a_Mallomonas_rasilis				-----				-----GTGAT	CGGACACGTC	GATGCTGGTA	AATCAACCAC	[35]
041_EF1a_Phaeodactylum_tricornutum				-----A	TGGGAAAAGA	A---AAGGTT	CACATTAGTT	TGGTCGTGAT	CGGACACGTC	GATGCTGGAA	AATCCACCAC	[68]
042_EF1a_Thalassiosira_pseudonana				-----A	TGCTTAAGGA	T---AAAACA	CACATATCGC	TCGTCGTGAT	TGGCCACGTT	GATTGTGGCA	AAAGTACAAC	[68]
043_EF1a_Malawimonas_jakobiformis				-----								[0]
044_EF1a_Malawimonas_californiana				-----A	TGGGCAAGGA	G---AAGACT	CACGTTAAACA	TTGTGTCAT	TGGACACGTC	GATGTCGGCA	AGTCGACCAC	[68]
045_EF1a_Trimastix_pyriformis				-----						-----AC		[2]
046_EF1a_Dinenympha_exilis				-----						-----TCAACAAC		[8]
047_EF1a_Monocercomonoides_sp.				-----A	TGGGAAAAGA	G---AAAGCA	CATATCAATC	TTGTTGTTAT	CGGTCACGTC	GATGTGGGAA	AATCCACAAC	[68]
048_EF1a_Streblomastix_strix				-----						---GTCGGCA	AATCCACGAC	[17]
049_EF1a_Histomonas_meleagridis				-----A	TGGGTAAGGA	G---AAAGAA	CATATCAATC	TCGTCGTTAT	CGGACACGTC	GATGCTGGTA	AATCAACAAC	[68]
050_EF1a_Trichomonas_vaginalis				-----A	TGGGTAAGGA	G---AAAGAG	CATATCAATA	TTGTCGTTAT	CGGACACGTC	GATGCAGGTA	AGTCGACAAC	[68]
051_EF1a_Hexamita_inflata				-----						---AACGGTA	AGTCTACCCT	[17]
052_EF1a_Spironucleus_vortens				-----						---AACGGTA	AGTCACTCT	[17]
053_EF1a_Andalucia_incarcerata				-----						---TCTGGCA	AGTCTACCAC	[17]
054_EF1a_Seculamonas_ecuadoriensis				-----A	TGGGAAAAGA	G---AAGGCT	CAC TTGAACC	TCGTCGTGAT	TGGCCACGTC	GA CTCTGGCA	AGTCGACCAC	[68]
055_EF1a_Reclinomonas_americana				-----A	TGGGCAAGGA	G---AAGGCT	CAC TTGAACC	TGGTCGTGAT	CGGCCACGTC	GA CTCTGGCA	AGTCGACCAC	[68]
056_EF1a_Histiona_aroides				-----A	TGGGCAAGGA	G---AAGGCT	CAC TTGAACC	TGGTCGTGAT	TGGCCACGTC	GACTCCGGCA	AGTCTACTAC	[68]
057_EF1a_Jakoba_libera				-----A	TGGGAAAAGA	G---AAGCTT	CAC TTGAACC	TGGTCGTGAT	CGGCCACGTC	GATTCCGGCA	AGTCGACCAC	[68]
058_EF1a_Stachyamoeba_lipophora				-----A	TGGGAAAAGA	G---AAGGCT	CATATTAACC	TCGTCGTGAT	CGGTCACGTC	GATGCTGGCA	AGTCGACCAC	[68]
059_EF1a_Naegleria_gruberi				-----								[2]
060_EF1a_Naegleria_andersoni				-----						---GCTGGTA	AGTCGACCAC	[17]
061_EF1a_Acrasis_rosea				-----						---TCTGGAA	AGAGCACAAC	[17]
062_EF1a_Sawyeria_marylandensis				-----A	TGGTAAAAGA	T---AAAATT	CACATTAACCT	TAGTTGTTAT	TGGTCACGTT	GACGCAGGTA	AATCAACAAC	[68]
063_EF1a_Bodo_saltans				-----						---GGCA	AGTCCACTGC	[14]
064_EF1a_Rhynchobodo_sp.				-----								[2]
065_EF1a_Rhynchomonas_nasuta				-----	--GGGAAGGA	G---AAGGTG	CACATCAGTC	TGGTCGTGCT	CGGCCACGTC	GATGCCGGCA	AGAGCACCGC	[65]
066_EF1a_Dimastigella_trypaniformis				-----					-----GTC	GATGCTGGCA	AGTCGACCAC	[23]

[illegible]

	120	130	140	150	160	170	180	190	200	210	220	
001_aEF1a_Desulfurococcus_mobilis	GACAGGACAC	ATACTGTACC	GTCTAGGATA	CTTCGACGAG	AAGACAGTGA	AGATGATCGA	GGAAGAGTCC	AAGAAGATGG	GTAAGGAGAG	CTTCAAGTTC	GCGTGGCTAC	[181]
002_aEF1a_Sulfolobus_solfataricus	AGTGGGAAGA	CTACTAATGG	ATAGAGGTTT	CATAGATGAA	AAAACAGTGA	AGGAAGATCG	AGAAAGTGTCT	AAAAAATTAG	GTAAGAATTC	TGAAAGAGTT	GCATTCTTAC	[175]
003_aEF1a_Thermoplasma_acidophilum	GgTaAggAcGg	CtGcTgTaGg	AgCaCgGcGa	GaTtCcAgCa	CaTaTtAtAg	AaGaStAtAg	AaAgGaAgCa	GaGcAgAaGg	GtAaGgCgAg	CtTcGaAtTt	GcCtGgGtCa	[178]
004_aEF1a_Halobacterium_sp.	GGTCGGGCGC	CTCCTCTACG	AAACGGGGAG	CGTACCCGAG	CACGTGATCG	AACAGCACAA	AGAAGAGGCC	GAGGAGAAGG	GCAAGGGCGG	CTTCGAGTTC	GCCTACGTCa	[175]
005_aEF1a_Methanocaldococcus_jannaschii	AGTCGGTAGA	TATTATATAG	ACAGTGGAGC	TATCGACCAA	CAGTGTATTAG	AGAAGTTAAA	AAGAGAAGCT	CAAGAGAGAG	GTAAGAGCAG	ATTCGAGTTT	GCTTACGTCa	[178]
006_EF1a_Acanthamoeba_culbertsoni	CACTGGCCAC	TTGATCTACA	AGTGCGGTGG	TATCGACAAG	CGTACCATTG	AGAAGTTTCA	GACTGAGGCC	AAGGAGATGG	GCAAGGGCTC	TTTCAAGTAC	GCCTGGGTGC	[151]
007_EF1a_Neoparamoeba_branchipila	CACTGGTCAC	TTGATCTACA	AGTGCGGAGG	TATTGACAAG	CGTGCCATCG	AGAAGTTTCA	GAAAGAGGCC	GCCGATATGG	GTAAGGGTTC	CTTCAAGTAC	GCCTGGGTTT	[130]
008_EF1a_Planoprotostelium_aurantium	CACCGGTAC	TTGATCTACA	AGTGCGGAGG	TATCGACAAG	AGAACCATCG	AGAAGTTTCA	GAAGGAGGCC	AAGGAGATCG	GAAAGGCCCT	TTTCAAGTAC	GCCTGGGTGC	[127]
009_EF1a_Physarum_polycephalum	CACTGGACAC	TTGATCTACA	AGTGCGGTGG	TATCGACAAG	CGTACCATTG	AGAAGTTTCA	GAAGGAAGCC	GCTGAAATGG	GCAAGGGTTC	CTTCAAGTAC	GCGTGGGTGT	[127]
010_EF1a_Amoebidium_parasiticum	TACCGGTAC	CTCATCTACA	AGTGCGGTGG	TATTGACAAG	CGTACCATTG	AGAAGTTTCA	AAAGGAAGCC	CAGGAAATGG	GCAAGGGCTC	CTTCAAGTAC	GCATGGGTGC	[148]
011_EF1a_Capsaspora_owczaraki	GACTGGTCAC	CTCATCTACA	AGTGCGGTGG	TATCGACAAG	CGCACCATCG	AGAAGTTTCA	GAAGGAAGCT	CAGGAAATGG	GCAAGGGCTC	TTTCAAGTAC	GCCTGGGTGC	[178]
012_EF1a_Monosiga_ovata	CACCGGCCAC	TTAATCTACA	AGTGCGGTGG	TATCGACAAG	CGTACCATTG	AGAAGTTTCA	GAAGGAGGCC	GCCGAGATGG	GCAAGGGCTC	CTTCAAGTAC	GCCTGGGTTC	[151]
013_EF1a_Salpingoeca_amphoridium	CACTGGTCAC	CTCATCTACA	AGTGCGGTGG	CATTGACAAG	CGTACCATTG	AGAAGTTTCA	GAAGGAGGCT	GCTGAGATGG	GCAAGGGTTC	GTTCAGATAC	GCATGGGTGC	[145]
014_EF1a_Ceanorhabditis_elegans	AACTGGACAC	TTGATCTACA	AGTGCGGAGG	AATTGACAAG	AGAACCATCG	AGAAGTTTCA	GAAGGAGGCT	CAGGAGATGG	GTAAGGGATC	TTTCAAGTAC	GCTTGGGTGC	[178]
015_EF1a_Drosophila_melanogaster	CACCGGCAC	TTGATCTACA	AGTGCGGTGG	TATCGACAAG	CGTACCATTG	AGAAGTTTCA	GAAGGAGGCC	CAGGAGATGG	GAAAGGGATC	CTTCAAGTAC	GCCTGGGTTC	[178]
016_EF1a_Ornithorhynchus_anatinus	CACAGGGCAC	CTCATCTACA	AATGTGGGGG	TATCGACAAG	AGAACCATCG	AAAAGTTTGA	GAAGGAGGCC	GCTGAGATGG	GGAAGGGCTC	CTTTAAATAT	GCCTGGGTTC	[178]
017_EF1a_Xenopus_laëvis	AACCGGCCAT	CTCATCTACA	AATGCGGGGG	CATCGACAAA	AGAACCATCG	AGAAGTTTGA	GAAGGAAGCC	GCAGAGATGG	GTAAGGGCTC	CTTTAAATAT	GCCTGGGTTC	[178]
018_EF1a_Saccharomyces_cerevisiae	TACCGGTCAT	TTGATTTACA	AGTGCGGTGG	TATTGACAAG	AGAACCATCG	AAAAGTTTGA	AAAGGAAGCC	GCTGAATTAG	GTAAGGGTTC	TTTCAAGTAC	GCTTGGGTTC	[178]
019_EF1a_Neurospora_crassa	TACCGGTAC	TTGATCTACA	AGTGCGGTGG	TATCGACAAG	CGTACCATTG	AGAAGTTTCA	GAAGGAAGCC	GCTGAGCTCG	GTAAGGGTTC	CTTCAAGTAC	GCCTGGGTTC	[181]
020_EF1a_Cryptococcus_neoformans	CACCGGTAC	TTGATCTACA	AGTGCGGTGG	TATCGACAAG	CGAACCATTCG	AGAAGTTTGA	GAAGGAGGCT	CAAGAGCTCG	GAAAGTCTTC	TTTCAAGTAC	GCTTGGGTTC	[178]
021_EF1a_Schizophyllum_comune	CACCGGCCAT	CTCATCTACA	AGTGCGGTGG	TATCGACAAG	CGTACGATCG	AGAAGTTTGA	GAAGGAGGCC	GCTGAGCTTC	GCAAGGGTTC	CTTCAAGTAT	GCCTGGGTGC	[178]
022_EF1a_Monoblepharis_insignis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[64]
023_EF1a_Rhizophlyctis_harderi	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[64]
024_EF1a_Harpochytrium.sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[16]
025_EF1a_Absidia_glauca	CACTGGTCAT	TTGATCTACA	AGTGCGGTGG	TATTGACAAG	CGTACCATTG	AAAAGTTTGA	AAAGGAAGCT	GCTGAAGTTC	GTAAGGGTTC	CTTCAAGTAC	GCCTGGGTGT	[178]
026_EF1a_Smittium_simulii	TACTGGTCAC	TTGATCTACA	AGTGCGGTGG	TATTGATAAA	CGTACTATTG	AAAAGTTTGA	GAAGGAAGCT	AACGAGCTCG	GTAAGGGTTC	ATTCAGATAC	GCTTGGGTTC	[151]
027_EF1a_Chara_australis	CACCGGCCAT	CTCATCTACA	AGCTGGGAGG	CATTGACAAG	CGTGTGATCG	AGAGGTTTGA	GAAGGAGGCC	GCTGAAATGA	ACAAGAGGTC	CTTCAAGTAC	GCCTGGGTGT	[127]
028_EF1a_Physcomitrella_patens	CACCTGGGCAC	CTCATCTACA	AGTTGGGAGG	TATCGACAAG	CGTGTGATCG	AGCGTTTTCGA	GAAGGAGGCC	GCGGAGATGA	ACAACGCTTC	GTTCAGATAT	GCCTGGGTGC	[178]
029_EF1a_Acetabularia_acetabulum	CACAGGTAC	TTGATCTACA	AACATAGGCG	TATCGACAAG	CGTGTGATCG	AGAAATTTGA	AAAGGAAGCC	GCAGAGATGA	ACAAGAGGTC	TTTCAAATAT	GCATGGGTGT	[127]
030_EF1a_Chaetomorpha_coliformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
031_EF1a_Arabidopsis_thaliana	AACTGGTCAC	TTGATCTATA	AGCTTGGTGG	TATTGACAAG	CGTGTGATCG	AGAGGTTTGA	GAAGGAGGCT	GCTGAGATGA	ACAAGAGGTC	CTTCAAGTAC	GCATGGGTGT	[178]
032_EF1a_Oryza_sativa	CACGGGCCAT	CTGATCTACA	AGCTTGGAGG	TATCGACAAG	CGTGTGATCG	AGAGGTTTGA	GAAGGAAGCC	GCTGAGATGA	ACAAGAGGTC	CTTCAAGTAC	GCCTGGGTGC	[178]
033_EF1a_Cyanidioschyzon_merolae	GACGGGTAC	TTGATCTACA	AGTGCGGCGG	CATCGACAAG	GCAACGATCG	AAAAGTTTGA	GAAAGAGGCC	AACGAAATGG	GAAAAGGTTT	GTTCAGATAC	GCCTGGGTGC	[178]
034_EF1a_Porphyrha_yezoensis	CACCGGCCAT	CTGATCTACA	AGTGCGGGGG	TATCGAGAAG	CGCGCGATTG	AGAAGTTTGA	GAAGGAGGCT	GCTGAGATGG	GCAAGGGCTC	GTTCAGATAC	GCCTGGGTGC	[178]
035_EF1a_Plasmodium_falciparum	TACAGGTCAC	ATTATTTACA	AATTAGGAGG	TATGATAGAG	AGAACTATTG	AAAAGTTTGA	GAAAGAACTC	GCTGAAATGG	GTAAGGGTAG	TTTTAAATAT	GCATGGGTTC	[178]
036_EF1a_Cryptosporidium_parvum	CACTGGTCAC	TTGATTTACA	AGCTCGGTGG	TATTGATAAG	AGAAGTATCG	AAAAGTTTGA	GAAGGAATCT	TCCGAAATGG	GTAAGGGTTC	CTTCAAATAT	GCCTGGGTTC	[178]
037_EF1a_Tetrahymena_pyriformis	TACCGGTAC	TTGATCTACA	AGTGCGGTGG	TATCGATAAG	AGAGTATCG	AAAAGTTTGA	GAAGGAATCC	GCTGAAACAG	GTAAGGGTTC	TTTCAAGTAC	GCTTGGGTTC	[181]
038_EF1a_Stentor_coeruleus	CACCGGTAC	TTAATCTATA	AGTGCGGTGG	TATTGACAAG	AGAACCATTG	AAAAGTTTGA	GAAGGAAGCT	AGCGAAATGG	GCAAGTCTTC	CTTCAAGTAT	GCATGGGTGC	[133]
039_EF1a_Phytophthora_parasitica	CACTGGTCAC	TTGATCTATA	AGTGCGGTGG	TATCGACAAG	CGTACCATTG	AGAAGTTTGA	GAAGGAGGCT	GCCGAGCTCG	GCAAGACCTC	GTTCAGATAC	GCCTGGGTGC	[178]
040_EF1a_Mallomonas_rasilis	CACTGGACAC	TTGATTTACA	AATGCGGTGG	TATTGATAAG	CGTACTATTG	AGAAGTATGA	GAAGGAAGCT	GCCGAAATGG	GTAATCATC	CTTCAAGTAC	GCCTGGGTGC	[145]
041_EF1a_Phaeodactylum_tricornutum	CACAGGTAC	TTGATCTACA	AATGTGGAGG	AATCGATAAG	CGTACCATTG	AAAAGTTTGA	AAAGGAAGCC	GCCGAACTCG	GAAAGGGATC	CTTCAAGTAT	GCATGGGTTC	[178]
042_EF1a_Thalassiosira_pseudonana	AACCTGGTCAT	CTCATATACA	AACATAGGAG	CATTGACAAA	AGAACAATTG	AGAAGTATGA	GAAAGAGCT	AATGAGCTCG	GCAAGGGCTC	ATTCAAATAT	GCTTGGGTGT	[178]
043_EF1a_Malawimonas_jakobiformis	-ACTGGTCAC	TTGATCTACA	AGTGCGGTGG	TATTGACAAG	CGCACCATCG	AGAAGTTTGA	GAAGGAAGCC	GCCGAGATGG	GCAAGGGCTC	CTTCAAGTAC	GCCTGGGTGC	[109]
044_EF1a_Malawimonas_californiana	CACCGGTAC	CTCATCTACA	AGTGCGGTGG	CATTGACAAG	CGCACCATCG	AGAAGTTTGA	GAAGGAGGCT	GCCGAGATGG	GTAAGGGCTC	GTTCAGATAC	GCCTGGGTGC	[178]
045_EF1a_Trimastix_pyriformis	CACTGGTCAC	CTGATCTACA	AGTGCGGTGG	TATTGATGAG	CGTACCATTG	AGAAGTTTGA	GGCTGAGGCC	AACGCCATCG	GCAAGGGTTC	CTTCAAGTAC	GCCTGGGTTC	[112]
046_EF1a_Dinenympha_exilis	GACGGGGCAC	TTGATCTACA	AGTGCGGGGG	CATTGATGAG	AGGACCATTA	AAAAGTTTGA	GACGAGGCTC	GAGGCCATGG	GAAAGGGTTC	GTTCAGATAT	GCATGGGTGC	[118]
047_EF1a_Monocercomonoides.sp.	AACCTGGACAT	CTGATCTACA	AGTGCGGTGG	TATTGATAAG	CGTACAATTG	AAAATTTTGA	GCAAGAAGCC	GATCAAAATCG	GAAAGGCTTC	ATTCAAATAT	GCCTGGGTGC	[178]
048_EF1a_Streblospio_saxicola	AACAGGACAT	CTGATCTACA	AGTGCGGAGG	AACCGACAAA	GCAACAATCG	AAAATTTTGA	GCAGGAGGCT	GAAACAGATTG	GGAAGGGCAT	TTTCAAGTAT	GCATGGGTAT	[127]
049_EF1a_Histomonas_meleagridis	AACAGGCCAC	TTAATCTACA	AGTGCGGTGG	TATCGATAAG	AGAAAGCTCG	CCCAACTTGA	AAAGGAAGCC	GTCCAATTC	GTAAGGGCTC	CTTCAAATAT	GCCTTCGTTA	[178]
050_EF1a_Trichomonas_vaginalis	AACAGGCCAC	CTCATCTACA	AGTGCGGGCG	TCTCGATAAG	CGTAAACTTG	CTGCCATCGA	AAAGGAAGCT	GAAACAATCG	GCAAGTCTTC	CTTCAAGTAC	GCTTTCGTTA	[178]
051_EF1a_Hexamita_inflata	CACTGGTCAC	TTGATCTACA	AGTGCGGTGG	TATCGATCAA	CGTACCTTGA	AGAAGTTTGA	AAAGGAAGCC	AACGAAATCG	GTAAGGGCTC	CTTCAAGTAT	GCCTGGGTTC	[127]
052_EF1a_Spiroplasma_vortens	CACTGGCCAC	TTGATCTTCA	AGTGCGGTGG	TATCGACCAG	CGTACCTTGA	AGAAGTTTGA	GAAGGAAGCC	AACGAGCTCG	GCAAGGGCTC	CTTCAAGTAT	GCCTGGGTTC	[127]
053_EF1a_Andalucia_incarcerata	CACTGGTCAT	CTCATCTACA	AGTGCGGTGG	CATCGACAAG	CGTACCATTG	AGAAGTTTGA	GAAGGAAGCC	GCTGAAATGG	GCAAGGGTTC	TTTCAAGTAT	GCCTGGGTTC	[127]
054_EF1a_Seculamonas_ecuadoriensis	CACTGGTCAC	CTGATCTACC	AGTGCGGGCG	TATCGACAAG	CGTACGATTG	AGAAGTTTGA	GAAGGAGGCT	AACGACATTG	GCAAGGGCTC	GTTCAGATAC	GCCTGGGTGC	[178]
055_EF1a_Reclinomonas_americana	CACTGGTCAC	CTCATCTACA	AGTGCGGTGG	CATTGACAAG	CGCACCATCG	AGAAGTTTGA	GAAGGAGGCC	AATGAGATGG	GCAAGGGCTC	GTTCAGATAC	GCCTGGGTGC	[178]
056_EF1a_Histiona_aroidea	CACTGGTCAC	TTGATCTACC	AATGCGGTGG	TATCGACAAG	CGTACCATTG	AGAAGTTTGA	GAAGGAAGCT	GCTGAAATGG	GCAAGGGCTC	TTTCAAGTAT	GCCTGGGTGT	[178]
057_EF1a_Jakoba_libera	CACCGGCAC	TTGATCTACC	AGTGCGGTGG	TATCGACAAG	CGAACCATTCG	AGAAGTTTGA	GAAGGAGGCC	TCCGAGATGG	GCAAGGGATC	CTTCAAGTAC	GCCTGGGTGC	[178]
058_EF1a_Stachyamoeba_lipophora	CACTGGTCAT	TTGATCTACA	AGTTGGGGCG	TATTGACAAG	CGTGTCTATTG	AGAAGTTTGA	GAAGGAGGCC	GCCGACATCG	GCAAGGGTTC	TTTCAAGTAT	GCCTGGGTTC	[178]
059_EF1a_Naegleria_gruberii	TACTGGTCAC	TTGATCTACA	AGTTGGGTGG	TATCGATAAG	AGAGTCATTG	AAAAGTTTGA	AAAGGAAGCT	GCTGAAATGG	GTAATCTCT	ATTCAAATAT	GCTTGGGTTC	[112]
060_EF1a_Naegleria_andersoni	CACTGGTCAC	TTGATCTACA	AGTGCGGTGG	TATCGATAAG	AGAGTCATTG	AAAAGTTTGA	AAAGGAAGCA	GCTGAAATGG	GTAATCTCT	ATTCAAATAT	GCTTGGGTTC	[127]
061_EF1a_Acrasis_rosea	CACTGGTCAC	TTGATCTACA	AGTGCGGAGG	TATTGACAAG	AGAGTCATTG	AGAAGTTTGA	GAAGGAAGCC	GCCGATATGG	GTAACAATAT	CTTCAAGTAC	GCCTGGGTTC	[127]
062_EF1a_Sawyeria_marylandensis	CACTGGTCAT	TTAATTTATA	AATGTGGTGG	AATTGACAAA	AGAGTCATTG	AAAATTTTGA	AAAAGAAGCA	CAAGAAATGG	GAAAAGGTTC	TTTCAAATAT	GCATGGGTTC	[178]
063_EF1a_Bodo_saltans	CACCGGTAC	TTGATCTACA	AGTGCGGTGG	TATTGACAAG	CGCACGATCG	AGAAGTTTGA	GAAGGAAGCC	GCCGAAATCG	GCAAGGGCTC	CTTCAAGTAT	GCCTGGGTTC	[124]
064_EF1a_Rhynchobodo.sp.	GACGGGTAC	TTGATCTACA	AGTGCGGTGG	TATTGACAAG	CGTACCATTG	AGAAGTTTGA	GAAGGAAGCC	GCCGAGATGG	GTAAGGGCTC	CTTCAAGTAC	GCTTGGGTGC	[112]
065_EF1a_Rhynchomonas_nasuta	CACCGGTAC	CTCATCTACA	AGTGCGGTGG	TATCGACAAG	CGCACCATTCG	AGAAGTTTGA	GAAGGAGGCC	GCAAGAAATGG	GCAAACTCTC	ATTCAAGTAT	GCCTGGGTGC	[175]
066_EF1a_Dimastigella_trypaniformis	CACCGGCCAC	CTGATCTACA	AGTGCGGGCG	CATTGACAAG	CGTACCATTG	AGAAGTTTGA	GAAGGAGGCC	GCCGAGATGG	GCAAGGGCTC	GTTCAGATAC	GCCTGGGTGC	[133]

	120	130	140	150	160	170	180	190	200	210	220	
133_EFL1a_Trachelomonas_volvocina	TACAGGCCAT	TTGATTTTCA	AGTCGCGAGG	TATTGACAAG	CGAACCATTTG	AGAAATTCGA	GAAAGGAAGCC	TCCGAAATGG	GCAAGGGTTC	ATTTCAAGTAC	GCGTGGGTGT	[175]
134_EFL_environmental_sample_EU810232	TACTGGACGT	CTCATCTTTG	AGCTGGGAGG	AATCCCCGAG	CGTGAGATGA	CCAAGCTTCG	TGCCGAGGCT	GAGGCTCTTG	GAAAGGGATC	TTTCGCCTTT	GCCTTCTACA	[178]
135_EFL_Creolimax_fragrantissima	----GGGCAT	CTAATCTTTG	CCCTCGGTGG	AATCCCCGAG	CGTGAGATGA	CCAAGCTTAA	GGAAGAGGCC	ACCCGTTTGG	GTAAGGGATC	TTTCGCCTTC	GCTTTCTATA	[106]
136_EFL_Sphaeroforma_arctica	TACCGGTTCAT	CTGATCTTCG	CCCTCGGGGG	TATCCCCGAG	CGTGAGATGC	AGAAGCTAAA	GGAAGAGGCC	ACCCGCTTGG	GTAAGGGTTC	GTTCGCCTTT	GCTTTCTACA	[178]
137_EFL_Monosiga_brevicollis	GACTGGCCGT	CTCATCTTCG	AGCTCGGGCG	CATCCCCGAG	CGTGAGATGC	AGAAGCTGAA	GGATGAGGCT	GAGCGTCTCG	GCAAGGGCTC	TTTCGCCTTC	GCCTTCTACA	[178]
138_EFL_Xenopus_laevis	CACCGGCCGT	CTGATCTTCG	AGCTGGGAGG	GATCCCCGAG	CGTGAGATGC	AGAAGCTCAA	GGATGAGGCC	CAGAAGCTCG	GCAAGGGCTC	GTTCGCCTTC	GCCTTCTACA	[178]
139_EFL_Spizellomyces_punctatus	CACCGGTCTG	CTTCTCTTCG	AGCTCGGTGG	TATCCCCGAG	CGTGAATGG	AGAAGCTTCG	TCAGGAGGCT	GAGCACCTCG	GCAAGTCTTC	CTTCGCCTTT	GCCTTCTACA	[178]
140_EFL_Gaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
141_EFL_Allomyces_macrogynus	GACGGGCCGT	CTGCTCTTCG	AGCTCGGGCG	TATCCCCGAG	CGTGAGATGG	AGAAGCTCCG	CCAGGAGGCC	GATGCCCTCG	GCAAGTCGTC	GTTCGCCTTC	GCCTTCTACA	[178]
142_EFL_Blastocladiella_emersonii	CACCGGTCTG	CTGCTCTTCG	AGCTCGGGCG	TATCCCCGAG	CGTGAGATGG	AGAAGCTCCG	CGCTGAGGCC	GATGCTCTCG	GCAAGTCGTC	GTTCGCCTTC	GCCTTCTACA	[178]
143_EFL_Catenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
144_EFL_Coelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
145_EFL_Entomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
146_EFL_Basidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
147_EFL_Conidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
148_EFL_Tetraselmis_tetrathele	CACCGGGCGC	CTGCTCTTCG	AGCTGGGGCG	TATCCCCGAG	CGTGAGATGG	ACAAGCTGAA	G6CCGAGGCC	GAGGCTCTCG	GCAAGCAGTC	CTTCGCCTTC	GCCTTCTACA	[127]
149_EFL_Mesostigma_viride	GACTGGCCGC	CTGATCTTCG	AGCTTGGAGG	TATCCCCGAG	CGTGATATGG	AGAAGCTGAG	G6CTGAGGCC	GACCGTCTCG	GAAAGTCTTC	CTTCGCCTTC	GCCTTCTACA	[181]
150_EFL_Ostreococcus_tauri	CACCTGGTCG	CTACTGTTTC	AGCTCGGGCG	TATCCCCGAG	CGTGAATCG	AAAAGCTCAA	G6CTGAGGCC	GATGCGCTCG	GCAAGTCCTC	TTTCGCCTTC	GCCTTCTACA	[178]
151_EFL_Nephroselmis_olivacea	CACAGGTCGC	CTTCTTTTCG	AGCTCGGGCG	TATCCCCGAG	CGTGAGATGG	AGAAGTTGAG	G6CTGAGGCC	GACCGCTTGG	GCAAGTCCTC	TTTCGCCTTC	GCCTTCTACA	[154]
152_EFL_Micromonas_pusilla	CACCGGGCGC	CTCATCTTCG	AGCTCGGAGG	CATCCCCGAG	CGCGAGATCG	AGAAGCTCCG	G6CGAGGCC	GATGCCCTCG	GCAAGTCCTC	CTTCGCCTTC	GCCTTCTACA	[127]
153_EFL_Parachlorella_kessleri	CACCGGGCGC	CTGCTGTTTC	AGCTGGGGCG	TATCCCCGAG	CGTGAGCTTG	AGAAGCTGAA	GGAAGAGGCC	CTGCGCTTTC	GCAAGTCCTC	CTTCGCCTTC	GCCTTCTACA	[136]
154_EFL_Helicospiridium_sp.	CACCGGTCTG	CTGCTCTTCG	AGCTCGGGCG	TATCCCCGAG	CGTGAGCTCG	AGAAGCTCAA	GGAAGAGGCC	GCTGCCCTGG	GCAAGTCCTC	CTTCGCCTTC	GCCTTCTACA	[178]
155_EFL_Goniium_pectorale	CACCTGGTCG	CTGCTGTTTC	AGCTCGGGCG	TATCCCCGAG	CGTGAGCTGG	AGAAGCTGAA	GGAAGAGGCC	GCTGCCCTGG	GCAAGTCGTC	GTTCGCCTTC	GCCTTCTACA	[178]
156_EFL_Chlamydomonas_reinhardtii	CACCTGGTCG	CTGCTGTTTC	AGCTCGGGCG	TATCCCCGAG	CGTGAGCTGG	AGAAGCTGAA	GGAAGAGGCC	GCTGCCCTGG	GCAAGTCCTC	TTTCGCCTTC	GCCTTCTACA	[178]
157_EFL_Scenedesmus_obliquus	CACCGGTCGC	CTGCTCTTCG	AGCTCGGGCG	TATCCCCGAG	CGTGAGCTGG	AGAAGCTCAA	GGAAGAGGCC	GCTGCCCTGG	GCAAGCAGTC	CTTCGCCTTC	GCCTTCTACA	[178]
158_EFL_Oriza_sativa	CACCGGTCGC	CTGCTCTTCG	AGCTCGGGCG	TATCCCCGAG	CGTGAGCTGG	AGAAGCTCAA	GGAAGAGGCC	GCTGCCCTGG	GCAAGCAGTC	CTTCGCCTTC	GCCTTCTACA	[178]
159_EFL_Chlorococcum_sp.	AACCGGGCGC	CTCATTTTTG	AGCTCGGGCG	TATTTCCGAG	CGTGAGGTTG	AGAAGTTGAA	GGAAGAGGCC	GCCACCTTCG	GTAAGTCCTC	CTTCGCCTTC	GCCTTCTACA	[127]
160_EFL_Ochlochaete_hystrix	CACCGGGCGC	CTGATATTCG	AGCTCGGGCG	TATCCCCGAG	CGTGAGTTGG	ACAAGCTGAA	GGAAGAGGCC	GCTGCCCTGG	GCAAGTCGTC	GTTCGCCTTC	GCCTTCTACA	[136]
161_EFL_Ulva_intestinalis	CACCGGGCGC	CTCATTTTTG	AGCTTGGCGG	CCTTCCCGAG	CGTGAGCTGG	AGAAGCTGAA	GGAAGAGGCC	GCCCATTTGG	GCAAGGCTTC	TTTCGCCTTC	GCCTTCTACA	[127]
162_EFL_Urospora_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
163_EFL_Gymnochlorella_stellata_ppc	GACTGGAAGG	CTTTTGTTCG	AACCTGGGCG	GATAGAAGAG	AGAGAGATGG	AGAGATGAA	ATCTGAAGCC	AAAGCTCTGA	ATAAGGAAAC	CTTTCTCTTT	GCTTTCTATA	[214]
164_EFL_Lotharella_amoebiformis_ppc	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
165_EFL_Chlorarachnion_reptans_ppc	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
166_EFL_Chondrus_crispus	CACCTGGCCGA	TTGATCTTCG	AGCTTGGAGG	CATCTCCGAG	CGTGAATGCG	AGAAGTTGAA	GGATGAGGCC	AAGGCCCTCG	GGAAGGATTC	CTTCGCATTC	GCCTTCTACA	[178]
167_EFL_Gracilaria_changii	CACCTGGCCGA	CTCATCTTCG	AGCTCGGGCG	CATTTCCGAT	CGTGAAATGC	AGAGCTCTTG	GAGGCTGATC	AAAGGCTTTC	GAAAGGATTC	TTTCGCCTTC	GCCTTCTACA	[178]
168_EFL_Cryptomonas_ovata	GACTGGTCGT	TTGCTGTTTC	AGCTCGGTGG	TATTTCCGAG	CGTGAGCTCG	AGAAGCTCAA	GGAAGAGGCC	GCCAACCTTC	GCAAGTCGTC	CTTCGCCTTC	GCCTTCTACA	[133]
169_EFL_Guillardia_theta	CACCTGGAGCT	CTCCTCTTCG	AGCTTGGAGG	TATCCCCGAG	CGTGAGCTCG	AGAAGCTCAA	GGAAGAGGCC	GCCAACCTTC	GCAAGTCCTC	CTTCGCCTTC	GCCTTCTACA	[175]
170_EFL_Rhodomonas_salina	CACCGGTCTG	CTTCTGTTTC	AGCTCGGGCG	TATCCCCGAG	CGTGAGCTTG	AGAAGCTCAA	GGAAGAGGCC	GCCAACCTTC	GCAAGTCCTC	CTTCGCCTTC	GCCTTCTACA	[175]
171_EFL_Pavlova_lutheri	GACCGGGCGC	CTGCTCTTCG	AGCTTGGCGG	CATCCCCGAG	CGCGAGATGG	AGAAGCTCCG	GCCCGAGGCC	GACCGCTTGG	GCAAGTCGTC	GTTCGCCTTC	GCCTTCTACA	[175]
172_EFL_Chrysochromulina_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
173_EFL_Isochrysis_galbana	CACCGGGCGC	CTTCTCTTCG	AGCTCGGGCG	TATTTCCGAG	CGTGAGATGG	AGAAGCTGAA	G6CCGAAGCC	GATCGCTTCG	GCAAGTCGTC	CTTCGCCTTC	GCCTTCTACA	[178]
174_EFL_Emiliania_huxleyi	CACCGGGCGG	CTGCTCTTCG	AGATGGGGCG	TATCCCCGAG	CGTGAGATGG	AGAAGCTCAA	G6CCGAGGCC	GACCGCTTGG	GCAAGTCGTC	CTTCGCCTTC	GCCTTCTACA	[175]
175_EFL_Achnanthes_kuwaitensis	CTGCGGTTCAC	TTGATCTTCA	AGCAGGGCGG	TATCTCCGAG	CGTGAGATGA	CCAAGCTCCA	G6CTATGGCC	GAGGAGAAGG	GTAAGTCCTC	CTTCGCCTTC	GCCTACTACC	[112]
176_EFL_Asterionella_glacialis	ATGCGGTTCAC	TTGATTTTCA	AGTGGCGTGG	TATCTCTGAG	CGTGAGATGA	AGAAGCTTCA	AGAAACTGCT	GAAGAGAAGG	GTAAGGCTTC	TTTCGGTTTC	GCTTACTACC	[112]
177_EFL_Detionula_confervacea	CTGCGGGCAT	CTTATCTTCA	G6CAGGGAGG	AATCTCTGCT	CGTGAGATGG	AAAAGCTCAA	G6CAACCGCC	GCCGAAAAGG	GTAAGTCCTC	TTTCGGATTTC	GCCTACTACC	[112]
178_EFL_Thalassiosira_pseudonana	CTGTGGACAC	TTGATCTTCA	AGCAGGGAGG	AATCTCCGAG	CGTGAGATGG	ACAAGCTTCA	G6CCACCGCC	GAAGAGAAGG	GAAAGTCCTC	TTTCGGATTTC	GCCTACTACC	[175]
179_EFL_Perkinsus_marinus	TACTGGTCGT	CTTCTTTTTC	AGCTGGGAGG	CATTCTGTAA	CGTGAATCGC	AGAAGCTCAA	GCAAGAGGCC	GATCGTCTTC	GTAAGTCGTC	TTTCGCCTTC	GCCTTCTACA	[187]
180_EFL_Oxyrrhis_marina	GACAGGCCAT	TTGTTGTTTC	AGTTGGGTGG	TCTTCCGGAG	CGTGAGCTGG	ACAAGCTTCG	CCAGGAGGCC	GATCGTTTTC	GAAGAGCAG	CTTCGCCTTC	GCTTTCTTCA	[184]
181_EFL_Karlodinium_micrum	AACAGGGCCG	CTGATCTTCG	AATTAAGTGG	GTTCGCCAGAA	GCGAGGCTTG	AAAAGCTGAA	GCAAGAGGCC	GAGCGTTTTC	GAAGAGGATC	TTTTCGCTTC	GCCTTTTACA	[184]
182_EFL_Amphidinium_carterae	GACCGGGCCG	TTGCTCTTTG	AGCTTGGTGG	TATCCCCGAG	CGTGAGTTGG	AGAAGTTGAA	G6CAGAGGCC	GAGCGCTTTC	GCAAGAGGCTC	GTTCGCCTTC	GCAATTCTACA	[184]
183_EFL_Heterocapsa_triquetra	GACTGGCAGG	CTGATTTTTG	AGCTCGGGCG	CCTGCGGGAG	CGTGAGCTGG	AGAAGCTGAA	GCAAGAGGCC	GAGCGCTTCG	GCAAGGGGCTC	TTTCGCCTTC	GCCTTCTACA	[184]
184_EFL_Planoglabrattella_operculis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[13]
185_EFL_Reticulomyxa_filosa	-----CGT	TATTTATTCG	AATTAAGGAG	TATGAGTGAC	AGAGAATTGG	AAAAATGAA	GAAAGAAGCT	CAAGAATTGG	GAAGAGATTTC	ATTTTGTGTT	GCTTTCTTAT	[103]
186_EFL_Lotharella_amoebiformis	YACCGGWCAC	CTTCTYTTCA	AGCTCGGAAC	WATGGACGAG	CGTGAGCGCA	AGGCGCTCAT	CGCTGAGGCC	GAGAAGGAGG	GCAAGGGTTC	ATTCGCCTTC	GCCTTCTTCA	[220]
187_EFL_Gymnochlorella_stellata	TACTGGTCAC	CTTCTTTTTC	AACTCGGTAC	TATGGACGAG	CGTATGAAAG	CTGACCTTAT	CAACAAAGGCA	AAGGAGATGA	AGAAGGAGTC	TTTCGCATTC	GCTTTCTTTC	[136]
188_EFL_Chlorarachnion_reptans	-----CAC	CTTCTCTTCG	AGCTCGGTAC	TATGGACGAC	CGTGAGAGG	CTGAGCTTTT	GGCCAAAGGCT	GAAGAGATGA	AGAAGAGGTC	TTTCGCCTTC	GCCTTCTTCA	[103]
189_EFL_Diplonema_ambulator	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
190_EFL_Diplonema_papillatum	GACTGGCCGT	TTGTTGTTTC	AGCTGGGAGG	AATCCCCGAG	CGTGAGATGG	AGAAGCTGCG	ACAGGAGGCC	GAGCGTCTGG	GCAAGCAGTC	GTTCGCCTTC	GCCTTCTACA	[151]
191_EFL_Rhynchopus_euleeides	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
192_EFL_Petalomonas_cantuscygni	-----	-----	-----	-----	-----	-----	-----	-----	-----TC	GTTCGCCTTC	TTGTTTCTACA	[22]
193_EFL_Trypanoplasma_borrel	RACCGGACGC	CTGCTRTACG	AYCTGGGWWG	TATTGACGAG	CGTGCCATGG	ARAAGCTRCG	KAAGATCGCT	GAGGAGAATG	GAAAGCCGTC	CTTTGCCTTC	GCACTACTWCA	[184]
194_EFL_Neobodo_salien	CACCGGTCTG	CTCTGTACG	ATATGGGGCG	CATYGACGAG	CGCACCATGG	AGAAGCTGAA	GAAGAAAGGCC	GAGGAGAAGC	GCAAGCCGTC	GTTCGCCTTC	GCCTACTACC	[133]

	230	240	250	260	270	280	290	300	310	320	330	
001_aEF1a_Desulfurococcus_mobilis	TCGACAGAACT	GAAGGAGGAG	CGTGAGCGTG	GAGTCACTAT	ATCGCTGTGG	TACATGAAGT	TCGAGACGAA	GAAGTACTTC	TTTACCATTCA	TCGATGCCCC	TGGCCACAGG	[291]
002_aEF1a_Sulfolobus_solfataricus	TAGATAAGT	AAAGAAGAG	AGAGAAGAG	GTGTAAACAT	TTTCACTAAG	TTTCACTAAG	TTTCACTAAG	TTTCACTAAG	TTTCACTAAG	TTTCACTAAG	TTTCACTAAG	[285]
003_aEF1a_Thermoplasma_acidophilum	TgGaTaGgTt	CaAgGaAgAg	AgGgAgAgAg	GaGtTaCcAt	TgAtCtGgCc	CaCaGgAaAt	TcGaGgAcCg	TaAgTaTtAc	TtCaCgCtTa	TcGaTgCgCc	TgGcCaCaGa	[288]
004_aEF1a_Halobacterium_sp.	TGGACAACCT	CGCGCAGGAG	CGCGAACCGG	GTGTCAACAT	CGACATCGCC	CACCAAGAGT	TCACCAACCG	CGAGTACGAG	TTTCACTATC	TCGATGTCTC	TGGCCACCGG	[285]
005_aEF1a_Methanocaldococcus_jannaschii	TGGACAACCT	GAAGAAGAG	AGAGAAGAG	GGGTTTCAAT	TGACGTAGCT	CACAAGAAAT	TCGAACCCCA	AAATATGAA	GTTACAATCG	TCGATGTCTC	AGGACACAGG	[288]
006_EF1a_Acanthamoeba_culbertsoni	TCGACAAGCT	CAAGGCTGAG	CGTGAGCGTG	GTATCACCAT	CGATATCGCT	CTGTGGAAGT	TCGAGACTGC	CAAGTACTAC	TTTACCATTCA	TTGATGCCCC	CGGTCAACCGT	[261]
007_EF1a_Neoparamoeba_branchipila	TGGACAAGCT	CAAGGCCGAG	CGTGAGAGAG	GTATCACCAT	TGATATTGCT	CTCTGGAAGT	TCGAGACTGC	CAAGTTCTAC	TTTACCATTCA	TCGATGCCCC	TGGTCAACCGT	[240]
008_EF1a_Planoprotostelium_aurantium	TCGACAAGCT	GAAGGCTGAG	AGAGAGCGTG	GAATCACCAT	CGACATTGCC	CTCTGGAAGT	TCGAGACCAC	CAAGTACTAC	TTTACCATTCA	TCGACGCCCC	AGGTCAACCGT	[237]
009_EF1a_Physarum_polycephalum	TGGACAAACT	CAAGTCCGAG	CGTGAGCGTG	GTATCACCAT	TGACATTGCT	CTCTGGAAGT	TCGAAACCGG	TAAGTACTAC	ATCACCATTCA	TCGATGCCCC	CGGACATCGT	[237]
010_EF1a_Amoebidium_parasiticum	TTGACAAACT	GAAGGCCGAG	CGCGAGCGCG	GTATCACCAT	CGACATCGCT	CTGTGGAAGT	TCGAGTCCAC	CAAGTACCAG	GTCCACCATCA	TCGATGCCCC	GGGCCACCGT	[258]
011_EF1a_Capsaspora_owczaraki	TCGACAAGCT	CAAGGCTGAG	CGCGAGCGTG	GCATCACCAT	CGACATTGCC	CTCTGGAAGT	TCGAGACCTC	CAAGTTCTAC	GTCCACCATCA	TCGATGCCCC	CGGCCACCGT	[288]
012_EF1a_Monosiga_ovata	TGGACAAGCT	GAAGGCCGAG	CGTGAGCGCG	GTATCACCAT	CGACATTGCT	CTGTGGAAGT	TCGAGACCAC	CAAGTTCTAT	GTCCACCATCA	TCGATGCCCC	CGGCCACCGT	[261]
013_EF1a_Salpingoeca_amphoridium	TTGACAAGCT	CAAGGCCGAG	CGCGAGCGCG	GTATCACCAT	CGACATTGCC	CTGTGGAAGT	TCGAGACTCC	CAAGTTCTAT	GTCCACCATCA	TCGATGCCCC	TGGCCACCGT	[255]
014_EF1a_Ceanorhabditis_elegans	TCGACAAGCT	GAAGGCCGAG	CGTGAACTGT	GTATCACCAT	CGATATCGCT	CTCTGGAAGT	TCGAGACTGC	CAAGTACTAC	ATCACCATTCA	TCGATGCCCC	AGGACATCGT	[288]
015_EF1a_Drosophila_melanogaster	TGGATAAGTT	GAAGGCTGAG	CGCGAGCGTG	GTATCACCAT	CGATATCGCC	CTGTGGAAGT	TCGAAACTGC	CAAGTACTAC	GTGACCATTCA	TTGATGCCCC	CGGACACAGG	[288]
016_EF1a_Ornithorhynchus_anatinus	TGGACAAGCT	GAAGGCTGAA	CGTGAACGAG	GCATCACCAT	TGACATTGCC	CTCTGGAAGT	TTGAGACGAC	CAAACTACTAC	ATCACCATTCA	TTGATGCCCC	TGGACACAGA	[288]
017_EF1a_Xenopus_laevis	TGGATAAGTT	GAAGGCTGAA	CGGGAACGTG	GCATCACCAT	CGATATTGCC	TTGTGGAAGT	TTGAGACTGG	AAAGTATTAT	ATCACCATTCA	TTGATGCTCC	TGGCCACAGA	[288]
018_EF1a_Saccharomyces_cerevisiae	TGGACAAGTT	AAAGGCTGAA	AGAGAAAGAG	GTATCACCAT	CGATATTGCT	TTGTGGAAGT	TCGAAACTCC	AAAGTACCAA	GTTCACGTTA	TTGATGCTCC	AGGTACACAGA	[288]
019_EF1a_Neurospora_crassa	TTGACAAGCT	CAAGGCCGAG	CGTGAGCGTG	GTATCACCAT	CGATATCGCT	CTGTGGAAGT	TCGAGACTCC	CAAGTACTAC	GTCCACGTTA	TCGATGCCCC	CGGTCACTCGT	[291]
020_EF1a_Cryptococcus_neoformans	TTGACAAGCT	TAAGGCCGAG	CGAGAGCGAG	GTATCACCAT	CGACATTGCC	CTTTGGAAGT	TCGAGACCCC	CAGGTACCAG	GTCCACGTTA	TTGACGCCCC	CGGTCAACCGA	[288]
021_EF1a_Schizophyllum_comune	TTGACAAGCT	CAAGGCTGAG	CGTGAGCGTG	GTATCACCAT	TGATATTGCT	CTGTGGAAGT	TCGAGACCCC	GAAGTACATG	GTCCACGTTA	TCGACGCCCC	GGGTCAACCGT	[288]
022_EF1a_Monoblepharis_insignis	TTGACAAGCT	CAAGGCCGAG	CGTGAGCGTG	GTATCACCAT	CGACATTGCC	CTGTGCGCTC	TCGAGACCCC	CAAGTACAA	GTTCACGTTA	TTGACGCCCC	TGGTCAACCGT	[174]
023_EF1a_Rhizophyctis_harderi	TTGACAAGTT	GAAGGCTGAA	CGTGAACGTG	GTATCACCAT	TGATATTGCT	CTTTGGAAGT	TCGAAACTCC	AAAGTACTAC	GTTCACGTTA	TTGATGCCCC	TGGACATCGT	[174]
024_EF1a_Harpochytrium_sp.	TGGACAAACT	GAAGGCTGAG	CGCGAGCGTG	GTATCACCAT	CGACATTGCC	CTCTGGAAGT	TCGAGACCCC	TAAGTCTTAC	GTCCACGTTA	TTGACGCCCC	GGGTCAACCGT	[126]
025_EF1a_Absidia_glauca	TGGACAAGCT	CAAGGCCGAG	CGTGAACGTG	GTATCACCAT	CGATATCGCT	CTCTGGAAGT	TCGAAACCCC	CAAGTACCAC	GTCCACGTTA	TTGACGCCCC	TGGCATCTGT	[288]
026_EF1a_Smittium_simulii	TGGACAAGCT	TAAGGCTGAA	CGTGAACTGT	GTATCACCAT	CGATATTGCT	CTGTGGAAGT	TCGAAACTCC	TAAGTACCTC	GTAACTGTTA	TTGATGCTCC	AGGTACACAGA	[261]
027_EF1a_Chara_australis	TAGACAAGCT	GAAGGCCGAG	CGCGAGCGTG	GAATTACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGACCAC	TAGGTATTAC	GTGACTGTCA	TCGATGCCCC	CGGCCATCGC	[237]
028_EF1a_Physcomitrella_patens	TGGACAAGCT	GAAGGCCGAG	CGCGAGCGTG	GTATCACCAT	TGATATTGCT	CTGTGGAAGT	TCGAGACTGC	GAAGTACTAC	TGCACGTGTA	TGCAGGCCCC	GGGACATCGC	[288]
029_EF1a_Acetabularia_acetabulum	TGGACAAGCT	CAAGGCCGAG	AGAGAGCGAG	GAATCACCAT	CGATATTGCC	CTTTGGAAGT	TTGAAACCCC	CAAGTACCAC	TGTACCGTTA	TTGACGCCCC	AGGACACAGG	[237]
030_EF1a_Chaetomorpha_coliformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[87]
031_EF1a_Arabidopsis_thaliana	TGGACAAACT	TAAGGCCGAG	CGTGAGCGTG	GTATCACCAT	CGATATTGCC	CTATGGAAGT	TCGAGACCAC	CAAGTACTAC	TGCACGTTA	TTGATGCCCC	AGGACATCGT	[288]
032_EF1a_Oryza_sativa	TCGACAAGCT	CAAGGCCGAG	CGTGAAGAGG	GTATCACCAT	CGATATCGCC	CTGTGGAAGT	TCGAGACCAC	CAAGTACTAC	TGCACGTTA	TCGATGCCCC	TGGTCAACCGT	[288]
033_EF1a_Cyanidioschyzon_merolae	TGGATAAACT	AAAGGCCGAA	CGGAGCGAG	GGATCACGAT	CGATATCGCG	CTGTGGAAGT	TCGAGACGGA	CAAGTACTAC	TTTACGATCA	TCGACGCTCC	TGGGCATCGT	[288]
034_EF1a_Porphyrha_yezoensis	TGGACAAACT	GAAGGCCGAG	CGCGAGCGCG	GTATCACCAT	TGACATTGCT	CTGTGGAAGT	TCGAGACGGA	GAAGTACAGC	TTTCACTATCA	TTGACGCCCC	GGGTCAACCGT	[288]
035_EF1a_Plasmodium_falciparum	TAGATAAACT	TAAGGCCAGAA	AGAGAAGAGG	GTATTACCAT	TGATATTGCC	TTTGAAGT	TTTGAACCCC	AGGTATTATC	TTTACTGTCA	TTGATGACCC	AGGTACACAG	[288]
036_EF1a_Cryptosporidium_parvum	TGGATAAGCT	CAAGGCTGAG	AGAGAAGAGG	GTATTACCAT	CGATATTGCC	TTTGAAGT	TCGAAACCCC	AAATACCAC	TACACTGTCA	TTGATGCCCC	AGGTACACAGA	[288]
037_EF1a_Tetrahymena_pyriformis	TGGACAAGCT	TAAGGCTGAA	AGAGAAGAGG	GTATCACCAT	CGATATTGCC	CTATGGAAGT	TCGAGACTGC	CAATACCAC	TTTACCATTCA	TCGATGCCCC	AGGACATAGA	[291]
038_EF1a_Stentor_coeruleus	TTGACAAGCT	TAAGGCTGAG	AGAGAAGCTG	GCATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAAACTGC	AAAGTCTTAT	TTTCACTATCA	TTGACGCCCC	TGGACACAGA	[243]
039_EF1a_Phytophthora_parasitica	TTGACAAGCT	GAAGGCCGAG	CGTGAGCGTG	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGTCCCC	CAAGTACTTC	TTTACGTTA	TTGATGCCCC	TGGTCAACCGT	[288]
040_EF1a_Mallomonas_rasilis	TTGATAACCT	CAAGGCCGAA	CGTGAACGCG	GTATCACCAT	TGATATTGCT	CTTTGGAAGT	TCGAGTCCCC	AAATTTTCA	TTTACTGTCA	TTGATGACCC	TGGACACAGA	[255]
041_EF1a_Phaeodactylum_tricornutum	TCGATAAACT	CAAGGCCGAA	CGTGAACGCG	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TTGAGTCTCC	CAAGTACTCT	TTTACCATTCA	TTGATGCCCC	TGGACATCGT	[288]
042_EF1a_Thalassiosira_pseudonana	TGGATCGTCT	CAAGGCCAGAA	CGTGAACGCG	GCATCACCAT	TGACATTGCT	CTTCAACGCT	TTGAGACTTC	CAATTTCAAG	TACACGTTA	TAGATGCCCC	AGGCCACAGA	[288]
043_EF1a_Malawimonas_jakobiformis	TTGACAAGCT	CAAGGCCGAG	CGCGAGCGCG	GTATCACCAT	CGATATTGCC	TTGTGGAAGT	TCGAGACCCC	CAAGTACAA	TTTACCATTCA	TTGATGCCCC	CGGACACCGC	[219]
044_EF1a_Malawimonas_californiana	TGGACAAGCT	CAAGGCCGAG	CGCGAGCGCG	GTATCACCAT	CGACATTGCC	CTGTGGAAGT	TTGAGACCCC	GAAGTACTAC	TTTACCATTCA	TCGATGCCCC	TGGTCAACCGT	[288]
045_EF1a_Trimastix_pyriformis	TCGACAAGTT	GAAGGCCGAG	CGTGAGCGTG	GTATCACCAT	TGATATTGCC	CTCTGGAAGT	TCGAGACTGA	GAAGTACTAC	TTTACCATTCA	TTGATGCCCC	CGGACACCGT	[222]
046_EF1a_Dinenympha_exilis	TGGACAAGTT	GAAGGCTGAG	CGCGAGCGTG	GTATCACCAT	TGATATTGCT	CTGTGGAAGT	TCGAGACCAC	CAAACTACTAC	TTTACGATTA	TTGATGCCCC	CGGACACAGG	[228]
047_EF1a_Monoceromonoides_sp.	TCGATAAACT	GAAGGCTGAG	CGTGAACGCG	GTATCACCAT	TGATATTGCC	TTGTGGAAGT	TCGAAACCAA	CAAACTACTAC	TTTACCATTCA	TTGATGCCCC	AGGACATAGA	[288]
048_EF1a_Streblospio_trix	TAGATAAGTT	GAAGGCCGAG	CGTGAGCGAG	GAATCACAAT	CGATATTGCC	CTGTGGAAGT	TTGAGACGAA	TAATACTAT	TTTACGATTA	TTGACGCCCC	TGGACACAGA	[237]
049_EF1a_Histomonas_meleagridis	TGGATAACCT	CAAGGCCGAA	CGTGAACGCG	GTATCACAAT	CGATATTGCC	CTTTGGAAGT	TCGAAATCAC	AAATACATG	TTTACAATCA	TTGATGCTCC	AGGACACAGA	[288]
050_EF1a_Trichomonas_vaginalis	TGGATCCCTC	CAAGGCTGAA	CGTGAGCGTG	GTATCACAAT	CGATATTGCC	CTCTGGAAGT	TCGAAAGGCCA	GAAGTTCTCC	TTTACCATTCA	TCGATGCCCC	AGGACACCGT	[288]
051_EF1a_Hexamita_inflata	TCGATCAACT	CAAGGCCGAG	CGTGAGCGTG	GTATCACAAT	CAACATCGCC	CTCTGGAAGT	TCGAAACCAA	GAAGTTCTAT	TTTACCATTCA	TCGATGCCCC	AGGTACACAGA	[237]
052_EF1a_Spironucleus_vortens	TCGACGAGCT	CAAGGCCGAG	CGTGAGCGTG	GTATCACCAT	CAACATCGCC	CTCTGGAAGT	TCGAGACCAC	GAAGTTCTAT	GTCCACATTCA	TCGACGCTCC	GGGCCACCGT	[237]
053_EF1a_Andalucia_incarcerata	TGGACAAGCT	GAAGGCTGAG	CGTGAACGCG	GTATCACCAT	CGACATCGCC	CTCTGGAAGT	TCGAGACCAC	GAAGTACTAC	TTTACCATTCA	TCGATGCCCC	AGGACATCGT	[237]
054_EF1a_Seculamonas_ecuadoriensis	TCGACAAGCT	GAAGGCCGAG	CGCGAGCGCG	GTATCACCAT	CGATATCGCC	CTGTGGAAGT	TCGAGACGCG	GAAGTACGTC	TGCACATTCA	TTGATGCCCC	GGGCCACCGT	[288]
055_EF1a_Reclinomonas_americana	TCGACAAGCT	GAAGGCCGAG	CGCGAGCGCG	GCATCACCAT	CGACATTGCC	CTCTGGAAGT	TCGAGACTCC	CAAGTACGTC	TGCACATTCA	TTGACGCCCC	CGGCCACCGT	[288]
056_EF1a_Histonia_aroidea	TGGACAAGTT	GAAGGCCGAG	CGCGAGCGCG	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGACTCC	CAAGTACGTC	TGCACATTCA	TTGATGCCCC	AGGACATCGT	[288]
057_EF1a_Jakoba_libera	TGGACAAGCT	CAAGGCCGAG	CGCGAGCGTG	GTATCACCAT	CGACATTGCC	TTGTGGAAGT	TCGAGACCCC	CAAGTACGTC	TGCACATTCA	TCGACGCCCC	TGGACACCGT	[288]
058_EF1a_Stachyamoeba_lipophora	TCGACAAGCT	GAAGGCCGAG	CGTGAGCGTG	GTATCACCAT	TGATATTGCT	CTGTGGAAGT	TCGAGACTCC	CAAGTACGAG	TTTACCATTCA	TTTACCATTCA	AGGACATCGT	[288]
059_EF1a_Naegleria_gruberi	TGGACAAGTT	GAAGGCTGAA	AGAGAAGAGG	GTATCACCAT	TGATATTGCT	CTGTGGAAGT	TTGAATCCAA	GAAGTACGTC	TTTACCATTCA	TTGATGCCCC	AGGTACACAGA	[222]
060_EF1a_Naegleria_andersoni	TGGACAAGTT	GAAGGCTGAA	AGAGAAGAGG	GTATCACCAT	TGATATTGCT	TTGTGGAAGT	TCGAAATCCAA	GAAGTACGTC	TTTACCATTCA	TTGATGCCCC	TGGACACAGA	[222]
061_EF1a_Acrasis_rosea	TGGACAAGCT	CAAGGCCGAA	CGTGAAGAGG	GTATCACCAT	CGATATTGCC	TTGTGGAAGT	TCGAGACCAC	CAAGACCATG	TTTACCATTCA	TTGATGCCCC	CGGACACAGA	[237]
062_EF1a_Sawyeria_marylandensis	TAGATAAACT	TAAAGCAGAA	AGAGAAGAGG	GTATTACTAT	TGATATTGCT	CTGTGGAAGT	TTGAACCCAA	AAATATTTC	TTTACCATTCA	TTGATGCCCC	AGGTACACAGA	[288]
063_EF1a_Bodo_saltans	TTGACAAGCT	GAAGGCTGAG	CGCGAGCGCG	GTATCACCAT	CGATATCGCC	CTCTGGAAGT	TCGAGTCCCC	CAAGTCCGTC	TTTACCATTCA	TTGATGCCCC	CGGACACCGT	[234]
064_EF1a_Rhynchobodo_sp.	TTGACAAGCT	GAAGGCTGAG	CGCGAGCGCG	GTATCACCAT	CGACATTGCC	CTCTGGAAGT	TCGAGTCCCC	CAAGTCCGTC	TTTACCATTCA	TCGATGCCCC	CGGACACCGT	[222]
065_EF1a_Rhynchomonas_nasuta	TTGACAAGCT	GAAGGCTGAG	CGCGAGCGCG	GTATCACCAT	CGACATCGCT	CTCTGGAAGT	TCGAGTCCCC	CAAGTCCGTC	TTTACCATTCA	TCGATGCCCC	CGGACACCGT	[285]
066_EF1a_Dimastigella_trypaniformis	TGGACAAGCT	GAAGGCTGAG	CGCGAGCGCG	GTATCACCAT	CGATATCGCT	CTGTGGAAGT	TCGAGTCCCC	CAAGTCCGTC	TTTACCATTCA	TCGACGCCCC	GGGCCACCGT	[243]

	230	240	250	260	270	280	290	300	310	320	330	
067_EF1a_Perkinsiella_amoebae	TCGACAAGCT	CAAGGCCGAG	CGCGAGCGCG	GTATCACGAT	CGACATCGCC	CTCTGGAAGT	TCGAGTCGCC	AAAGTCCGTC	TTACAGATCA	TCGACGCCCC	GGGGACCCGT	[213]
068_EF1a_Leishmania_major	TCGACAAGCT	CAAGGCCGAG	CGCGAGCGCG	GCATCACGAT	CGACATTGCG	CTGTGGAAGT	TCGAGTCGCC	CAAGTCCGTC	TTACAGATCA	TCGATCGCCC	CGGCCACCGC	[288]
069_EF1a_Leishmania_infantum	TCGACAAGCT	GAAGGCCGAG	CGCGAGCGCG	GCATCACGAT	CGACATTGCG	CTGTGGAAGT	TCGAGTCGCC	CAAGTCCGTC	TTACAGATCA	TCGATCGCCC	CGGCCACCGC	[288]
070_EF1a_Leishmania_braziliensis	TCGACAAGCT	GAAGGCCGAG	CGCGAGCGCG	GCATCACGAT	CGACATTGCG	CTGTGGAAGT	TCGAGTCGCC	CAAGTCCGTC	TTACAGATCA	TCGATCGCCC	CGGCCACCGC	[288]
071_EF1a_Leishmania_tarentolae	TGGACAAGCT	GAAGGCCGAG	CGCGAGCGCG	GCATCACGAT	CGACATTGCG	CTGTGGAAGT	TCGAGTCGCC	CAAGTCCGTC	TTACAGATCA	TCGATCGCCC	CGGCCACCGC	[288]
072_EF1a "L. tarentolae" FJ807257	TGGACCGTCT	GAAGGCCGAG	CGTGAAGCGC	GTATCACGAT	CGATATTGCG	CTGTGGAAGT	TCGAGTCCCC	GAAGTCCGTC	TTACAGATCA	TCGACGCGCC	TGGCCACCGT	[213]
073_EF1a_Leptomonas_costaricensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
074_EF1a_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
075_EF1a_Leptomonas_podlipaevi	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
076_EF1a_Herpetomonas_muscarum	TCGACAAGCT	GAAGGCCGAG	CGCGAGCGCG	GTATCACGAT	CGATATCTCC	CTCTGGAAGT	TCGAGTCCCC	GAAGTCCGTC	TTACAGATCA	TCGATCGCCC	CGGCCACCGT	[201]
077_EF1a_Herpetomonas_pessoai	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
078_EF1a_Blastocrithidia_culicis	TGGACAAGCT	GAAGGCCGAG	CGCGAACGCG	GTATCACGAT	CGATATTGCC	CTCTGGAAGT	TCGAGTCCCC	CAAGTCCGTC	TTACCATCA	TCGATCGCCC	CGGCCACCGC	[69]
079_EF1a_Trypanosoma_brucei	TGGACAAGCT	GAAGGCCGAG	CGCGAACGTC	GTATCACGAT	CGACATTGCA	CTGTGGAAGT	TCGAGTCACC	CAAGTCTGTC	TTCACTATTA	TTGATGCTCC	TGGGCCACCGT	[288]
080_EF1a_T. brucei_equiperdum	TGGACAAGCT	GAAGGCCGAG	CGCGAACGTC	GTATCACGAT	CGACATTGCA	CTGTGGAAGT	TCGAGTCACC	CAAGTCTGTC	TTCACTATTA	TTGATGCTCC	TGGGCCACCGT	[210]
081_EF1a_Trypanosoma_aviium	TCGACAAGCT	GAAGGCCGAG	CGTGAAGCGC	GTATCACGAT	CGACATTGCG	CTGTGGAAGT	TCGAGTCACC	CAAGTCTGTC	TTCACTATTA	TTGATGCTCC	TGGGCCACCGT	[234]
082_EF1a_Trypanosoma_cruzi	TGGACAAGCT	GAAGGCCGAG	CGCGAGCGCG	GTATCACGAT	CGACATTGCG	CTGTGGAAGT	TCGAGTCGCC	CAAGTCTGTC	TTACAGATCA	TCGACGCGCC	CGGCCACCGC	[288]
083_EF1a_Trypanosoma_rangeli	TGGACAAGCT	GAAGGCCGAG	CGCGAGCGCG	GTATCACGAT	CGACATTGCG	CTCTGGAAGT	TCGAGTCGCC	CAAGTCCGTC	TTACAGATCA	TCGATGCTCC	CGGCCACCGT	[288]
084_EF1a_Ploetitia_costatum	TCGACAAGCT	GAAGGCCGAG	CGTGAACGTC	GTATCACGAT	TGATATCGCT	CTCTGGAAGT	TCGAGACTGC	CAAGTCTGTT	TTACCATCA	TTGATGCAAC	CGGCCATCGT	[285]
085_EF1a_Ploetitia_oblonga_cDNA	TCGACAAGCT	GAAGGCCGAG	CGCGAGCGTC	GTATCACGAT	TGATATCGCT	CTCTGGAAGT	TCGAGACTGC	CAAGTCTGTT	TTACCATCA	TTGATGCAAC	CGGCCATCGT	[285]
086_EF1a_Ploetitia_oblonga_gDNA	TAGACAAGCT	CAAGGCCGAG	CGCGAGCGTC	GTATCACGAT	CGACATTGCC	TTGTGGAAGT	TCGAGACGCG	CAAGTCCGTC	TTACAGATCA	TTGATGCCCC	CGGACACCGT	[285]
087_EF1a_Entosiphon_sp.	TCGACAAGCT	GAAGGCCGAA	CGCGAGCGTC	GTATCACCAT	CGACATTGCC	CTGTGGAAGT	TCGAGACCCC	GAAGTCTGTC	TTACCATCA	TTGATGCCCC	GGGACATCGT	[285]
088_EF1a_Entosiphon_sulcatum	TCGACAAGCT	GAAGGCCGAG	CGCGAGCGTC	GTATCACCAT	CGACATTGCC	CTGTGGAAGT	TTGAGACCCC	CAAGTCCGTC	TTACCATCA	TTGATGCCCC	GGGACACCGT	[213]
089_EF1a_Peranema_tricophorum	TGGACAAACT	CAAGGCAGAA	CGTGAAGAG	GTATCACCAT	CGACATCGCC	TTGTGGAAGT	TCGAAACAGC	CAAACTCAGT	TTACCATCA	TTGATGCCCC	AGGACACAGA	[285]
090_EF1a_P. tricophorum_FJ807243	TGGACAAACT	YAAGGCCGAA	AGAGAAAGAG	GTATCACCAT	CGACATTGCC	CTTTGGAAGT	TCGAAACTGC	CAAGTCAGTC	TTACCATCA	TTGATGCTCC	AGGACATCGT	[279]
091_EF1a_Colacium_elongatum	TGGACAAGCT	TAAGGCTGAG	CGTGAAGCGTC	GTATCACCAT	CGATATTGCC	CTCTGGAAGT	TCGAGACTGC	AAAGTCTGTC	TTACCATCA	TCGATGCCCC	TGGCCATCGT	[285]
092_EF1a_Colacium_mucronatum	TGGACAAGCT	CAAGGCTGAG	CGTGAACGTC	GTATCACCAT	CGATATTGCC	CTCTGGAAGT	TCGAGACGCG	CAAGTCCGTC	TTACCAATCA	TCGATGCCCC	TGGTCAACCG	[231]
093_EF1a_Colacium_sp.	TGGACAAGCT	CAAGGCTGAG	CGTGAACGTC	GTATCACCAT	CGACATTGCC	CTCTGGAAGT	TCGAGACGCG	AAAGTCTGTC	TTACCATCA	TCGATGCCCC	TGGCCATCGT	[285]
094_EF1a_Colacium_vesiculosum	TGGACAAGCT	CAAGGCTGAG	CGTGAACGTC	GTATCACCAT	TGATATCGCT	CTCTGGAAGT	TCGAGACGCG	CAAGTCCGTC	TTACCAATCA	TCGATGCCCC	TGGTCAACCG	[243]
095_EF1a_Cryptoglena_pigra	TGGACAAACT	CAAGGCTGAG	CGTGAACGTC	GTATCACCAT	CGATATTGCC	CTCTGGAAGT	TCGAGACTGC	AAAGTCAAGTC	TTACCATCA	TTGACGCCCC	TGGCCATCGT	[285]
096_EF1a_Cryptoglena_skujai	TGGACAAGCT	CAAGGCCGAG	CGTGAAGCGTC	GTATCACGAT	TGATATTGCC	CTGTGGAAGT	TCGAGACTGC	GAAGTCAAGTC	TTACCATCA	TTGACGCCCC	TGGCCATCGT	[285]
097_EF1a_Discoplistis_adunca	TGGACAAGCT	CAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	CGATATTGCC	CTCTGGAAGT	TCGAGACGCG	CAAGTCTGTC	TTACCGTGA	TTGATGCCCC	CGGCCATCGT	[285]
098_EF1a_Euglena_clara	TGGATAAGCT	GAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	CGACATCGCC	CTGTGGAAGT	TCGAACTGCG	CAAGTCCGTC	TTACCATCA	TCGATGCCCC	CGGCCACCGC	[165]
099_EF1a_Euglena_deses	TGGATAAGCT	MAAGGCCGAG	CGTGAACGTC	GTATCACCAT	CGACATTGCC	CTGTGGAAGT	TCGAAACCGC	CAAGTCCGTC	TTACCAATTA	TTGATGCAAC	AGGCCATCGT	[285]
100_EF1a_Euglena_exilis	TGGACAAGCT	AAAGGCCGAG	CGTGAACGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAAACCGC	CAAGTCTGTC	TTACCATCA	TTGATGCCCC	TGGCCATCGT	[165]
101_EF1a_Euglena_geniculata	TCGACAAGCT	GAAGGCCGAG	CGTGAACGTC	GTATCACCAT	TGATATCGCT	CTCTGGAAGT	TCGAGACTGC	CAAGTCTGTT	TTACCATCA	TTGATGCAAC	CGGCCATCGT	[285]
102_EF1a_Euglena_gracilis	TGGACAAGCT	TAAGGCCGAG	CGTGAACGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAACTGCG	TAAGTCCGTC	TTACCATCA	TTGATGCCCC	TGGACACCGT	[285]
103_EF1a_E. gracilis_X16890	TGGACAAGCT	CAAGGCCGAG	CGTGAAGCGTT	GTATCACGAT	TGATATTGCC	CTGTGGAAGT	TCGAGACTGC	GAAGTCAAGTC	TTACCATCA	TTGATGCTCC	AGGACATCGT	[288]
104_EF1a_E. gracilis_FJ807238	TGGACAAGCT	TAAGGCCGAG	CGTGAACGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAACTGCG	TAAGTCCGTC	TTACCATCA	TTGATGCCCC	TGGACACCGT	[288]
105_EF1a_Euglena_granulata	TTGACAAGCT	TAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAAACAGC	TAAGTCTGTC	TTCACTATCA	TTGATGCTCC	TGGCCATCGT	[285]
106_EF1a_Euglena_hiemalis	TGGACAAGCT	TAAGGCCGAG	CGTGAACGTC	GTATCACCAT	TGATATTGCC	CTGTGGAAGT	TCGAACTGCG	CAGATCCGTC	TTACCATCA	TTGATGCAAC	TGGCCATCGT	[285]
107_EF1a_Euglena_longa	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
108_EF1a_Euglena_stellata	TGGATAAGCT	GAAGGCCGAG	CGTGAACGTC	GTATTTACCAT	TGATATTGCA	CTTTGGAAGT	TTGAACTGCG	GAAGTCCGTC	TTCACTATCA	TTGACGCCCC	AGGACATCGT	[165]
109_EF1a_Euglena_viridis	TGGACAAGCT	TAAGGCCGAA	CGTGAACGTC	GTATTTACCAT	TGATATTGCT	CTGTGGAAGT	TCGAGACGCG	CAAGTCTGTC	TTCACTATCA	TCGATGCCCC	TGGCCATCGT	[243]
110_EF1a_Eutreptia_pertyi	TTGACAAGCT	TAAGGCCGAG	CGTGAAGCGTC	GTATCACAT	CGATATTGCC	CTCTGGAAGT	TCGAGACGCG	TAAGTCTGTC	TTCACTATCA	TTGATGCTCC	TGGCCATCGT	[285]
111_EF1a_Eutreptiella_pomquetensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
112_EF1a_Lepocinlis_acus_var._major	TGGACAAGCT	CAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	TGATATTGCC	CTGTGGAAGT	TCGAGACTGC	GAAGTCAAGTC	TTACCATCA	TTGATGCCCC	TGGACATCGT	[285]
113_EF1a_Lepocinlis_ovum	TGGACAAGCT	CAAGGCCGAA	CGTGAAGCGTC	GTATCACCAT	TGATATTGCC	TTGTGGAAGT	TCGAACTGCG	CAAGTCCGTC	TTACCGTTA	TCGATGCCCC	TGGCCATCGT	[165]
114_EF1a_Lepocinlis_spyrogyroides_clone_38	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
115_EF1a_L. spyrogyroides_clones_1_&_4	TTGACAAGCT	GAAGGCCGAG	CGTGAACGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGACAGC	GAAGTCTGTC	TTACCAATTA	TTGACGCTCC	TGGCCATCGT	[285]
116_EF1a_Lepocinlis_steinii	TGGACAAGCT	CAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	TGATATTGCC	CTGTGGAAGT	TCGAACTGCG	CAAGTCTGTC	TTACCAATCA	TTGATGCAAC	TGGCCATCGT	[165]
117_EF1a_Lepocinlis_tripteris	TGGACAAGCT	CAAGGCCGAG	CGTGAACGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGACTGC	AAAGTCAAGTC	TTACAGTCA	TTGACGCCCC	TGGCCATCGT	[285]
118_EF1a_Monomorphina_aenigmatica	TGGACAAGCT	CAAGGCCGAG	CGCGAGCGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGACAGC	CAAGTCCGTC	TTACCAATTA	TCGATGCTCC	TGGGCCACCG	[285]
119_EF1a_Monomorphina_pyrum	TTGACAAGCT	GAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGACAGC	GAAGTCTGTC	TTACCAATTA	TTGACGCTCC	TGGCCATCGT	[285]
120_EF1a_Monomorphina_sp.	TTGACAAGCT	GAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	TGATATTGCC	CTGTGGAAGT	TCGAGACGCG	GAAGTCCGTC	TTACCATCA	TTGACGCTCC	TGGCCATCGT	[285]
121_EF1a_Phacus_orbicularis	TGGACAAGCT	CAAGGCCGAA	CGTGAAGCGTC	GTATCACCAT	TGACATTGCC	CTGTGGAAGT	TCGAGACTGC	CAAGTCAAGTC	TTACCGTCA	TTGACGCTCC	CGGCCATCGT	[243]
122_EF1a_Phacus_similis	TCGACAAGCT	TAAGGCCGAG	CGCGAGCGTC	GTATCACGAT	TGATATTGCC	CTGTGGAAGT	TCGAGACGCG	GAAGTCCGTC	TTACCATTA	TCGATGCAAC	CGGCCATCGC	[285]
123_EF1a_Strombomonas_acuminata	TGGACAAGCT	TAAGGCCGAG	CGTGAACGTC	GTATCACCAT	CGACATTGCC	TTGTGGAAGT	TTGAGACTGC	CAAGTCAAGTC	TTACCATTA	TTGATGCTCC	TGGCCATCGT	[288]
124_EF1a_Strombomonas_borsteniensis	TGGACAAGCT	TAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGACTGC	AAAGTCTGTC	TTACCATCA	TTGATGCCCC	TGGCCATCGT	[165]
125_EF1a_Strombomonas_napiformis	TGGACAAGCT	TAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	TGATATTGCC	CTGTGGAAGT	TTGAGACTGC	AAAGTCTGTC	TTACCATCA	TCGATGCCCC	CGGCCATCGT	[285]
126_EF1a_Strombomonas_ovalis	TGGACAAGCT	TAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	YGATATTGCC	CTGTGGAAGT	TCGAGACTGC	CAAGTCTGTC	TTACCATCA	TTGATGCCCC	CGGCCATCGT	[285]
127_EF1a_Strombomonas_triquetra	TGGACAAGCT	TAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	CGATATTGCC	CTGTGGAAGT	TCGAGACTGC	CAAGTCTGTC	TTACCATCA	TTGATGCCCC	TGGTCAACCG	[285]
128_EF1a_Trachelomonas_conradi	TGGACAAGCT	CAAGGCCGAG	CGTGAAGCGTC	GTATCACGAT	CGATATTGCC	TTGTGGAAGT	TCGAGACTGC	GAAGTCAAGTC	TTACCAATCA	TCGATGCCCC	AGGACATCGT	[285]
129_EF1a_Trachelomonas_echinata	TGGACAAGCT	TAAGGCCGAG	CGTGAACGTC	GTATTTACCAT	CGATATTGCC	CTTTGGAAGT	TCGAACTGCG	AAAGTCTGTT	TTACCATCA	TCGACGCCCC	CGGACATCG	[285]
130_EF1a_Trachelomonas_ilefvelei	TGGACAAGCT	CAAGGCCGAG	CGAGAGCGTC	GTATCACCAT	TGATATTGCC	CTGTGGAAGT	TCGAGACGCG	GAGGTCAAGTC	TTACCAATCA	TCGACGCCCC	TGGCCATCGT	[285]
131_EF1a_Trachelomonas_plancvonia	TGGACAAGCT	CAAGGCCGAG	CGAGAACGTC	GTATCACCAT	CGACATTGCC	CTGTGGAAGT	TCGAGACAGC	GAGGTCAAGTC	TTACCAATCA	TCGACGCCCC	CGGCCATCGT	[165]
132_EF1a_Trachelomonas_similis	TGGACAAGCT	CAAGGCCGAG	CGTGAAGCGTC	GTATCACCAT	TGATATTGCT	CTCTGGAAGT	TCGAACTGCG	CAAGTCAAGTC	TTACCATCA	TTGATGCCCC	TGGACACCGA	[165]

	230	240	250	260	270	280	290	300	310	320	330	
133_EFL1a_Trachelomonas_volvocina	TGGACAAGCT	CAAGGCCGAG	CGTGAGCGTG	GTATCACAAT	TGATATTGCT	CTGTGGAAGT	TCGAGACAGC	CAAGTCAGTC	TTCACGATCA	TTGATGCTCC	TGGTCACCGA	[285]
134_EFL_environmental_sample_EU810232	TGGACAGGCA	GAAGGACGAG	AGGGGSCGKG	GKGTGACCAT	CCAGTGCACC	CTGTTCACCA	CTGTTCACCA	CTGTTCACCA	TACACCATCA	TTGATGCCCC	GGGACACAGG	[288]
135_EFL_Creolimax_fragrantissima	TGGATCGTCA	GAAGGATGAG	CGTGAGCGTG	GTGTTACCAT	TGCTTGCACC	ACCAAGGAGT	TTTTCTACTCC	TACCAAGCAC	TACACTGTCA	TTGACGCCCC	TGGTCACAGA	[216]
136_EFL_Sphaeroforma_arctica	TGGATCGCCA	GAAGATGAG	CGTGAGCGTG	GTGTCAACAT	CGCTTGCACC	ACCAAGGAGT	TTTTCTACTCC	TACCAAGCAC	TACACCGTTA	TTGACGCCCC	CGGCCATAGA	[288]
137_EFL_Monosiga_brevicollis	TGGACCGTCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	TGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAACCAAGCAC	TACACCGTCA	TTGATGCCCC	TGGCCATCGT	[288]
138_EFL_Xenopus_laevis	TGGATCGCCA	GAAGGACGAG	CGTGAGCGTG	GTGTGACCAT	CGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAACCAAGCAC	TACACCGTCA	TCGACGCCCC	GGGCCATCGC	[288]
139_EFL_Spizellomyces_punctatus	TGGACCGTCA	AAAGGAGGAG	CGTGCCCGTG	GTGTGACCAT	TGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAACTACCAC	TACACCATCA	TTGACGCCCC	CGGTCAACCGT	[288]
140_EFL_Gaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
141_EFL_Allomyces_macrogynus	TGGACCGTCA	GAAGGAGGAG	CGCGCTCGTG	GTGTGACCAT	CGCGTGCACC	ACCAAGGAAT	TCTTCACTGC	CGGTGGGAC	TACACCATCA	TCGACGCCCC	CGGTCAACCG	[288]
142_EFL_Blastocladiella_emersonii	TGGATCGCCA	GAAGGAGGAG	CGCGCTCGTG	GTGTGACCAT	CGCGTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGAC	TACACCATCA	TCGACGCCCC	CGGCCACCGC	[288]
143_EFL_Catenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
144_EFL_Coelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
145_EFL_Entomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
146_EFL_Basidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
147_EFL_Conidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
148_EFL_Tetraselmis_tetrathele	TGGACAGGCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	CAAGTACCAC	TACACGATCA	TTGATGCCCC	CGGCCATCGC	[237]
149_EFL_Mesostigma_viride	TGGATCGCCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	TTCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAAGTGGGAC	TACACCATCA	TTGATGCCCC	CGGCCATCGC	[291]
150_EFL_Ostreococcus_tauri	TGGATCGCCA	AAAGGAAGAG	CGTGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TCGATGCCCC	GGGTACAGAG	[288]
151_EFL_Nephroselmis_olivacea	TGGATCGCCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TCGATGCCCC	CGGCCATCGC	[264]
152_EFL_Micromonas_pusilla	TGGACAGGCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	CAACTGGGAC	TACACCATCA	TCGACGCCCC	CGGTCAACCG	[237]
153_EFL_Parachlorella_kessleri	TGGACAGGCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	TGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAGGTGGGAC	TACACCATCA	TCGACGCCCC	GGGCCACAGG	[246]
154_EFL_Helicospiridium_sp.	TGGACAGGCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TTGATGCCCC	CGGCCATAGG	[288]
155_EFL_Goniomonas_pectoralis	TGGACCGGCG	CAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	TGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAAGTGGGAC	TACACCATCA	TCGATGCCCC	GGGTCAACCG	[288]
156_EFL_Chlamydomonas_reinhardtii	TGGACCGGCG	TAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	CGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CGGTGGGAC	TACACCATCA	TCGACGCCCC	CGGTCAACCG	[288]
157_EFL_Scenedesmus_obliquus	TGGACAGGCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TCGACGCCCC	AGGCCACAGG	[288]
158_EFL_“Oryza_sativa”	TGGACCGGCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	TGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAAGTGGGAC	TACACCATCA	TCGATGCCCC	AGGACACAGG	[288]
159_EFL_Chlorococcum_sp.	TGGACAGGCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	CAACTGGGAC	TACACCATCA	TTGATGCCCC	CGGCCATCGT	[237]
160_EFL_Ochlochaete_hystrix	TGGACAGGCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	TGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAGGTGGGAC	TACACCATCA	TCGATGCCCC	AGGCCACAGG	[246]
161_EFL_Ulva_intestinalis	TGGATCGCCA	AAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TTGATGCCCC	TGGCCATCGT	[237]
162_EFL_Urospora_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[90]
163_EFL_Gymnochlorella_stellata_ppc	TGGATCGGCA	AAAGGAAGAG	AGAGAGCGTG	GTGTTACGAT	CACATGCGCT	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TCGATGCCCC	TGGCCACAGG	[324]
164_EFL_Lotharella_amoebiformis_ppc	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[69]
165_EFL_Chlorarachnion_reptans_ppc	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[96]
166_EFL_Chondrus_crispus	TGGACAAGTC	CAAGGAGGAG	CGTGCCCGTG	GTGTGACCAT	CGCGTGCACC	ACCAAGGAGT	TCTTCACTGC	CTCTTACCAT	TACACGATCA	TTGATGCCCC	GGGCCATCGC	[288]
167_EFL_Gracilaria_changii	TGGACAAGTC	CAAGGAGGAG	TGCGCCGATG	GTGTGACCAT	TGCTTGCACC	ACCAAGGAGT	TTTTCACTGC	CTCTTATCAC	TACACCATCA	TTGATGCCCC	GGGCCATCGT	[288]
168_EFL_Cryptomonas_ovata	TGGACCGTCA	GAAGGAGGAG	CGTGAGCGTG	GTGTTACCAT	CGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TCGACGCTCC	GGGTACAGAG	[243]
169_EFL_Guillardia_theta	TGGACAGGCA	GAAGGAGGAG	CGTGAGCGTG	CGGTGACCAT	CGCATGCACT	ACCAAGGAGT	TCTTCACTGC	CAAGTGGGAC	TACACCATCA	TCGATGCCCC	GGGCCATAGG	[285]
170_EFL_Rhodomonas_salina	TGGACAGGCA	GAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	CGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	CAAGTGGGAC	TACACCATCA	TCGATGCCCC	GGGCCACAGG	[285]
171_EFL_Pavlova_lutheri	TGGACCGGCA	GAAGGAGGAG	CGCGAGCGTG	CGGTGACCAT	CTCGTGCACC	ACGAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACGATCA	TCGATGCCCC	CGGCCACCGC	[285]
172_EFL_Chrysochromulina_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[195]
173_EFL_Isochrysis_galbana	TGGACCGGCA	AAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TCGATGCCCC	GGGTCAACCG	[288]
174_EFL_Emilia_huxleyi	TGGACCGGCA	AAAGGAGGAG	CGCGAGCGTG	CGGTGACCAT	CTCGTGCACC	ACCAAGGAGT	TCTTCACTGC	CAAGTGGGAC	TACACCATCA	TCGATGCCCC	GGGCCACCGC	[285]
175_EFL_Achnanthes_kuwaitensis	TTGATACGTC	CAAGGAGGAG	CGCGAGCGTG	GTGTGACCAT	TCAGTGCAAC	ACCAAGGAGT	TCTTCACTGC	CAAGTACCAC	TACACTATTG	TCGATGCCCC	CGGCCACAGG	[222]
176_EFL_Asterionella_glacialis	TCGATACCTG	TAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	TCAGTGCAAC	ACCAAGGAGT	TTTTCACTGC	CAAGTACCAC	TACACTATTG	TCGATGCCCC	CGGCCACAGG	[222]
177_EFL_Detionula_confervacea	TCGACACCTG	TAAGGAGGAG	CGCGAGCGTG	GAGTTACCAT	TCAGTGTAAC	ACCAAGGAGT	TCTTCACTGC	CAAGTACCAC	TACACGATTC	TCGATGCCCC	CGGCCACAGG	[222]
178_EFL_Thalassiosira_pseudonana	TTGACACCTG	TAAGGAGGAG	CGTGAGCGTG	GAGTTACCAT	CCAGTGTAAC	ACCAAGGAGT	TCTTCACTGC	AAAGTACCAC	TACACCATTC	TCGATGCCCC	CGGCCACAGG	[285]
179_EFL_Perkinsus_marinus	TGGACCGGCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	CGCGTGCACC	ACGAAGGAGT	TCTTCACTGC	TCTGTGGGAC	TACACTATTA	TCGATGCCCC	TGGCCATAGG	[297]
180_EFL_Oxyrrhis_marina	TGGATCGGCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	CGCTTGCACC	ACGAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACTGTCA	TTGACGCCCC	TGGCCACCGC	[294]
181_EFL_Karlodinium_micrum	TGGACCGTCA	GAAGGAGGAG	CGCGAGCGTG	GAGTGACGAT	TGATGCACT	ACGAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACGGTGA	TCGATGCCCC	TGGCCATCGG	[294]
182_EFL_Amphidinium_carterae	TGGACAGGCA	GAAGGAGGAG	CGTGAGCGTG	GAGTGACGAT	TGCTTGCACC	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCATCA	TCGATGCCCC	TGGCCACAGG	[294]
183_EFL_Heterocapsa_triquetra	TGGACAGGCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACGAT	CGCATGCACT	ACCAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACTGTCA	TTGATGCCCC	GGGCCACCGT	[294]
184_EFL_Planoglabratella_operculis	TGGATAAGTC	AAAGGATGAG	AGAGTTCGCG	GTGTGACCAT	CCAGTGCTGC	ACTCGGAGT	TCTTCACTGC	CTCTTATCAT	TACACAATTA	TTGACGCCCC	GGGTCAACCG	[123]
185_EFL_Reticulomyxa_filosa	TGGACGATC	AAAGGAGGAG	AGAGCCCGTG	GTGTGACCAT	TGTTTGCAAT	ACCTGAGGAGT	TCTTCACTGC	AAAGTGGGAC	TACACAATTA	TTGATGCCCC	GGGTCAACCG	[213]
186_EFL_Lotharella_amoebiformis	TGGACAAGCA	AAAGGAAGAG	CGTGAGCGTG	GTGTGACCAT	CTCCTGCACC	ACCAAGGAGT	TCCACACAC	CAACTACCAC	TACACTGTCA	TCGACGCCCC	GGGACACCGT	[330]
187_EFL_Gymnochlorella_stellata	TGGACAAGCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	TTCTTGCACC	ACCAAGGAGT	TCCACACTAC	CAACTTCCAT	TACACGTC	TCGACGCTCC	TGGACACAG	[246]
188_EFL_Chlorarachnion_reptans	TGGATAAGCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	TGCTTGCACC	ACCAAGGAGT	TCCACACTAC	TAACTTCCAC	TACACCGTGA	TTGATGCCCC	TGGACATAG	[213]
189_EFL_Diplonema_ambulator	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
190_EFL_Diplonema_papillatum	TGGACAGGCA	GAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	CGCGTGCACC	ACGAAGGAGT	TCTTCACTGC	GAAGTGGGAC	TACACCGTGA	TCGACGCCCC	GGGACACCGT	[261]
191_EFL_Rhynchopus_euleeides	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
192_EFL_Petalomonas_cantuscygni	TGGATCGGTC	CAAGGAGGAG	CGTGAGCGTG	GTGTGACCAT	CCAGTGCACT	CTCAAGGAGT	TCTTCACTGC	CCAGTTCAC	TACACCATCA	TCGATGCCCC	GGGCCACCGC	[132]
193_EFL_Trypanoplasma_borreli	TGGACAACCA	GAAGGCCGAG	CGMGAGCGTG	GTATCACCAT	TGCTTGCACC	ACCAAGGAGT	TCTACACGCA	RCGCTACCAC	TACACGATCA	TYGACGCTCC	YGGCATCGT	[294]
194_EFL_Neobodo_salienis	TGGACAACCA	GAAGGCCGAG	CAGGAGCGTG	GTATCACCAT	TGCTTGCACC	ACCAAGGAGT	TCTACACGCA	CGCTACCAC	TACACCATCA	TCGACGCCCC	GGGCCACCGT	[243]

	340	350	360	370	380	390	400	410	420	430	440	
001_aEF1a_Desulfurococcus_mobilis	GACTTCGTTA	AGAACATGAT	TACCGGTGCA	AGCCAGGGTG	ATGCAAGCAT	CCTCGTG6TG	AGCGCTAGGA	AGGGTGAGTT	CGAGGCGAGG	ATGAGTGCTG	AG-----	[393]
002_aEF1a_Sulfolobus_solfataricus	GACTTCGTAA	AGAACATGAT	AATCGGTGCT	AGCCAAGCAG	ATGCGCGAAT	TCTAGTG6TT	TCAAGTGAAT	AGGGTGAATA	TGAGAGCTGA	ATGAGCTGAG	AA-----	[387]
003_aEF1a_Thermoplasma_acidophilum	GaCtTcGtCa	AgAaCaTgAt	AaCaGgTaCa	AgTcAgGcCg	AtGcGgCcAt	CcTaGtCaTa	TcTgCaAgGg	AaGgAgAaGg	TgTgAtG---	-----	-----	[375]
004_aEF1a_Halobacterium_sp.	GACTTCGTGA	AGAACATGAT	TACGGGTGCG	TCCCAAGGCC	ACACAGCCGT	CCTCGTGCTC	CGCGCTGACG	ACGGTGTCGC	G-----	-----	-----	[366]
005_aEF1a_Methanocaldococcus_jannaschii	GACTTCATTA	AAAACATGAT	TACAGGAGCT	TACACAGGAG	TACAGGAGCT	CTTAGTTGTT	GATGTTAAAT	ATGCCAAGAC	AGGAATTCA	-----	-----	[378]
006_EF1a_Acanthamoeba_culbertsoni	GATTTTCATCA	AGAACATGAT	CACCTGGTACC	TCTCAGGGTG	ACGTTGCCAT	CCTCGTCATT	GCCTCCGGTG	AGGGTGAGTT	CGAGGCCGGT	ATCTCCAAGA	AC-----	[363]
007_EF1a_Neoparamoeba_branchipila	GATTTTCATCA	AGAACATGAT	TACCGGTACC	TCTCAAGCGC	ATGTTGCCGT	CCTCGTCATT	GCCTCCGGTA	CTGGTGAGTT	CGAGGCTGGT	ATCGCCAAGA	AC-----	[342]
008_EF1a_Planoprotostelium_aurantium	GACTTCATCA	AGAACATGAT	TACCGGTACC	TCTCAGGGCG	ATTG6GTAT	CCTCGTCATT	GCCTCCCCCA	CCGAGAGATT	CGAGGCTGGT	ATCGACAAGA	AC-----	[339]
009_EF1a_Physarum_polycephalum	GATTTTCATCA	AGAACATGAT	CACCTGGTACT	TCCCAAGCCG	ATGCCGCCGT	GTTGGTGATT	GCCTCCCCCA	CTGGTGAGTT	CGAGGCTGGT	ATTGCCAAGA	CC-----	[339]
010_EF1a_Amoebidium_parasiticum	GATTTTCATCA	AGAACATGAT	CACGGGTACC	TCCCAAGCCG	ATTG6GTG	GCTCATCATC	GC08CCGGTA	TCGGTGAGTT	CGAGGCTGGT	ATCTCCAAGG	AT-----	[360]
011_EF1a_Capsaspora_owczaraki	GATTTTCATCA	AGAACATGAT	CACCGGTACC	TGCGAGGGTG	ACTGCGCCGT	CCTCATCGTT	GC08CCGGCA	CTGGTGAGTT	CGAGGCTGGC	ATCTCCAAGG	AC-----	[390]
012_EF1a_Monosiga_ovata	GACTTCATCA	AGAACATGAT	CACCGGCACC	TCCCAAGCCG	ACTGCGCCGT	CCTCATCGTC	GCTGCCGGCA	CTGGTGAGTT	CGAGGCTGGT	ATCTCCTCGA	AC-----	[363]
013_EF1a_Salpingoeca_amphoridium	GACTTCATCA	AGAACATGAT	CACGGGTACC	TGCGAGGGTG	ACTGCGCCGT	GCTCATCGTC	GCCGCCGGTA	CTGGTGAGTT	CGAGGCTGGT	ATCTCGTCCA	AC-----	[357]
014_EF1a_Ceanorhabditis_elegans	GATTTTCATCA	AGAACATGAT	CACCTGGAAACA	TCCCAAGCCG	ATTGCGCCGT	CTTGGTCGTC	GCTTGC6GAA	CTGGAGAATT	CGAGGCCGGA	ATCTCCAAGA	AC-----	[390]
015_EF1a_Drosophila_melanogaster	GATTTTCATCA	AGAACATGAT	CACCTGGTACC	TGCGAGGGCG	ATTGCGCCGT	CCTGATTGTC	GC08CCGGAA	CCGAGAGAAT	CGAGGCAAGT	ATCTCGAAGA	AC-----	[390]
016_EF1a_Ornithorhynchus_anatinus	GACTTCATCA	AGAACATGAT	TACTGGTACT	TCCCAAGCTG	ACTGTGCAAGT	CCTAATTGTG	GCTGCCGGCG	TGGGTGAGTT	CGAAGCTGGT	ATCTCCAAGA	AT-----	[390]
017_EF1a_Xenopus_laevis	GACTTCATCA	AAAACATGAT	TACTGGCACT	TCTCAGGGTG	ACTGTGCTGT	GCTCATGTTT	GCTGGTG6TG	TGGAGAATT	TGAGGCTGGT	ATCTCTAAAA	AT-----	[390]
018_EF1a_Saccharomyces_cerevisiae	GATTTTCATCA	AGAACATGAT	TACTGGTACT	TCTCAAGCTG	ACTGTGCTAT	CTTGATTATT	GCTGGTG6TG	TCGGTGAATT	CGAAGCCGGT	ATCTCTAAGG	AT-----	[390]
019_EF1a_Neuropora_crassa	GATTTTCATCA	AGAACATGAT	TACTGGTACC	TCCCAAGCTG	ACTGTGCTAT	CCTCATGATT	GC08CCGGTA	CTGGTGAGTT	CGAGGCTGGT	ATCTCCAAGG	AT-----	[393]
020_EF1a_Cryptococcus_neoformans	GACTTCATCA	AGAACATGAT	CACCGGTACC	TCCCAAGGGTG	ACTGTGCCAT	CCTCATGATT	GCCACCGGTA	TCGGTGAGTT	CGAGGCCGGT	ATCTCCAAGG	AC-----	[390]
021_EF1a_Schizophyllum_comune	GACTTCATCA	AGAACATGAT	CACCGGTACC	TCCCAAGGGTG	ACTGCGCTAT	CCTCACCATC	GC08GGTGGA	CTGGTGAATT	CGAGGCTGGT	ATCTCCAAGG	AT-----	[390]
022_EF1a_Monoblepharis_insignis	GACTTCATTA	AGAACATGAT	CACGGGTACC	TGCGAGGGTG	ACTGCGCCAT	CCTCATCATC	GCTGGTG6TA	CCGGTGAGTT	CGAGGCTGGT	ATCTCGAAGG	AC-----	[276]
023_EF1a_Rhizophlyctis_harderi	GATTTTCATCA	AGAACATGAT	TACTGGTACT	TCCCAAGCTG	ATTGTGCCAT	CTTGGTCATT	GCTTCCGGTG	TCGGTGAAAT	CGAAGCTGGT	ATCTCCAAGG	AC-----	[276]
024_EF1a_Harpochytrium_sp.	GATTTTCATCA	AGAACATGAT	CACCTGGTACT	TCCCAAGCCG	ACTGCGCCGT	CCTCATCGTC	GC08CCGGTA	CCGAGAGATT	CGAGGCCGGT	ATCTCCAAGG	AC-----	[228]
025_EF1a_Absidia_glauca	GATTTTCATCA	AGAACATGAT	TACTGGTACT	TCTCAAGCTG	ATTGCGGTA	CTTATCATCT	GCTGCGGTA	CTGGTGAATT	CGAAGCTGGT	ATCTCCAAGG	AT-----	[390]
026_EF1a_Smittium_simulii	GATTTTCATCA	AGAACATGAT	TACTGGTACT	TCTCAAGCTG	ATTGTGCCAT	TCTTATTATT	GCTGGTG6TG	TTGGTGAAAT	CGAAGCTGGT	ATCTCTAAGG	AC-----	[363]
027_EF1a_Chara_australis	GATTTTCATCA	AGAACATGAT	CACGGGTACC	TCCCAAGCCG	ACTGCGCTGT	GCTGGTCATC	GACTCGACGA	CGGGCGGGTT	CGAGGCCGGT	ATTTCAAAGG	AT-----	[339]
028_EF1a_Physcomitrella_patens	GATTTTCATCA	AGAACATGAT	CACGGGAACG	TGCGAGGCAG	ACTGCGCTGT	GCTGATCATC	GACTCGACGA	CGGAGGGTT	CGAGGCCGGT	ATCTCAAAGG	AT-----	[390]
029_EF1a_Acetabularia_acetabulum	GATTTTCATCA	AAAATATGAT	CACAGGAACG	TACAGGCAG	ACTGCGCTGT	GTTAATGATT	GACTCTACGC	CTGGTGAGTT	CGAGTCAGGC	ACCTCCAAGG	AC-----	[339]
030_EF1a_Chaetomorpha_coliformis	GATTTTCATCA	AGAATATGAT	TACTGGTACT	ACACAGGGTG	ACTGTGCTGT	TCTTATGATT	GACTCTACTC	CTGGTG6TTT	CGAATCAGGT	ATCTCCAAGG	AC-----	[189]
031_EF1a_Arabidopsis_thaliana	GATTTTCATCA	AGAACATGAT	TACTGGTACC	TCCCAAGCTG	ATTGTGCTGT	TCTTATGATT	GACTCCACCA	CTGGAAGTTT	TGAGGCTGGT	ATCTCTAAGG	AT-----	[390]
032_EF1a_Oryza_sativa	GACTTCATCA	AGAACATGAT	CACGGGTACC	TCCCAAGCTG	ACTGTGCCGT	GCTCATCACT	GACTCCACCA	CTGGTG6TTT	TGAGGCTGGT	ATCTCCAAGG	AT-----	[390]
033_EF1a_Cyanidioschyzon_merolae	GACTTTATCA	AGAACATGAT	CACGGGTACC	TCTCAGGGCG	ACCTTGCAT	CTTGGTG6TT	GC08CTCCTC	CTGGAAGATT	TGAGGCTGGT	ATCTCGTCGA	AC-----	[390]
034_EF1a_Porphyrha_yezoensis	GACTTCATCA	AGAACATGAT	CACGGGCACG	TGCGAGGGCG	ATGTGGCCAT	CCTGGTCATT	GCTTCCGCCG	CGGGCGAGTT	TGAGGCGGGT	ATCTCCCAAG	AC-----	[390]
035_EF1a_Plasmodium_falciparum	GATTTTATTA	AAAATATGAT	TACAGGTACA	TACAGAAGCA	ATGCTGCTTT	ATTAGTTGTC	CAAGCTGAAG	TAGGTTGTTT	TAGGTTGTTT	TTTTCAAAGG	AA-----	[390]
036_EF1a_Cryptosporidium_parvum	GATTTTCATCA	AGAATATGAT	TACTGGTACC	TCTCAAGCTG	ATGTTGCTTT	ATTGGTTGTC	CAAGCC-----	-GATCGTTT	CGAAGGTGCC	TTCTCCAAGG	AA-----	[384]
037_EF1a_Tetrahymena_pyriformis	GATTTTCATCA	AGAATATGAT	TACTGGTACC	TCCCAAGCTG	ATGTCGCTAT	TCTTATGATT	GCTTCCCCCT	AGGGTGAAAT	CGAAGCTGGT	ATCTCCAAGG	AC-----	[393]
038_EF1a_Stentor_coeruleus	GACTTTATCA	AGAACATGAT	TACAGGYACT	TACAGAAGCT	ATGTTGCCAT	CCTTATCGTC	CTTTCAGGAA	AAGGTGAATT	CGAAGCTGGC	ATTGCGAAGA	AC-----	[345]
039_EF1a_Phytophthora_parasitica	GACTTCATCA	AGAACATGAT	TACGGGGACC	TGCGAGGGCG	ATGCTGCCAT	CCTTGG6TG	GCCTCGGGTG	TGGGTGAGTT	CGAGGCTGGT	ATCTCCAAGG	AG-----	[390]
040_EF1a_Mallomonas_rasilis	GATTTTCATCA	AGAACATGAT	TACTGGAACA	TCTCAGGCCG	ACGTGGCTGT	CCTTGTATT	GATGGCGCAC	CTGGTG6TTT	CGAGGCTGGG	TGGGTTCTG	AT-----	[357]
041_EF1a_Phaeodactylum_tricornutum	GATTTTCATCA	AGAACATGAT	CACCGGAACT	TCCCAAGGGTG	ATGTTGCTGT	TCTCGTTATT	GACTGTGCC	AGGGAAGGTT	CGAAGCCGGT	ATCTCTAAGG	AC-----	[390]
042_EF1a_Thalassiosira_pseudonana	GATTTTATCA	AGAATATGAT	CACAGGGACG	TCTCAGGCCG	ACGTAGCTCT	TTTGGTGATT	GATGCCGCTC	AAGGAAAGTT	CGAGGCCGGT	ATATCTAAGG	AA-----	[390]
043_EF1a_Malawimonas_jakobiformis	GATTTTCATCA	AGAACATGAT	TACCGGTACC	TCCCAAGCCG	ACGCAAGCGT	GCTGGTTGTC	GCCGCCACCA	CTGGTGAGTT	CGAGGCTGGT	ATCTCCAAGG	AC-----	[321]
044_EF1a_Malawimonas_californiana	GATTTTCATCA	AGAACATGAT	CACCGGTACC	TCCCAAGGGTG	ATGCTGCCGT	GCTGTGCTTT	GCTGCCGGTA	CTGGTGAGTT	CGAGGCTGGT	ATCTCCAAGG	AC-----	[390]
045_EF1a_Trimastix_pyriformis	GACTTCATCA	AGAACATGAT	TACCGGTACC	TCTCAGGGCG	ATTGCGCTCT	CCTCGTCGTC	GCCGCCACCG	CTGGTGAGTT	CGAGGCCGGT	ATCTCCGGTG	AG-----	[324]
046_EF1a_Dinenympha_exilis	GATTTTCATCA	AGAACATGAT	TACGGGGACC	TGCGAGGGTG	ATGCTGCTAT	TCTGGTTGTT	GCTGCCAAGC	TTGGTGAGTT	TGAGGCTGGA	ATTTCAAAGG	AT-----	[330]
047_EF1a_Monocercomonoides_sp.	GATTTTCATCA	AGAACATGAT	CACCGGAACA	TACAGAAGCT	ATGTTGCTCT	TCTTGTGCTT	GCTGCCAAGC	TTGGTGAAAT	TGAAGCCGGT	ATTTCAAAGG	AT-----	[390]
048_EF1a_Streblospio_trix	GACTTCATCA	AGAACATGAT	AACAGGTACA	TCATAAGCTG	ATGCTGCATT	GCTTGTAGTA	GCTGCCAATT	AAGGAGAGTT	TGAAGCAGGT	ATCTCAAAAG	AC-----	[339]
049_EF1a_Histomonas_meleagridis	GACTTCATCA	AGAACATGAT	CACAGGTACA	TCCCAAGCCG	ATGCTGCCGT	CTTGTGTTAT	GATGCCACAC	GC08TG6TTT	CGAAGCCGGT	ATTGCGCAAC	AA-----	[390]
050_EF1a_Trichomonas_vaginalis	GACTTCATCA	AGAACATGAT	CACAGGCACA	TCCCAAGGGTG	ATGCTGCTAT	CTTGTGCTAT	GACTCCACAC	TCGGTG6CTT	CGAAGCTGGT	ATCGCTGAAC	AA-----	[390]
051_EF1a_Hexamita_inflata	GACTTCATCA	AGAACATGAT	CACCTGGTACC	TCCCAAGGGTG	ACGTGCGCAT	CCTCGTCGTC	GC08CCGGTG	AGGGTGAAAT	CGAAGCCGGT	ATCTCAAGCG	AA-----	[339]
052_EF1a_Spiroplasma_vortens	GACTTCATCA	AGAACATGAT	CACCGGTACC	TCCCAAGCCG	ACGTGCGCAT	CCTCGTCGTC	GC08CCGGTG	AGGGTGAGTT	CGAGGCTGGT	ATCTCCAAGG	AG-----	[339]
053_EF1a_Andalucia_incarcerata	GATTTTCATCA	AGAACATGAT	CACCTGGTACC	TCCCAAGGGTG	ATGCTGCCAT	TCTTGTGCTC	GCTTCCGGAA	CTGGTGAAAT	CGAAGCTGGT	ATCTCCAAGG	AA-----	[339]
054_EF1a_Seculamonas_ecuadoriensis	GACTTCATCA	AGAACATGAT	TACCGGTACC	TGCGAGGGCG	ATGCCGCCAT	CCTCGTCATT	GCCTCGGGCA	CTGGTGAGTT	CGAGGCCGGT	ATCTCGTCGG	AG-----	[390]
055_EF1a_Reclinomonas_americana	GACTTCATCA	AGAACATGAT	TACCGGCACG	TGCGAGGGTG	ATGCCGCCAT	TCTGATTGTT	GCTTCCGGCA	CTGGTGAGTT	CGAGGCTGGC	ATTCGCAAGG	AG-----	[390]
056_EF1a_Histiona_roides	GATTTTCATCA	AGAACATGAT	TACTGGTACT	TCTCAAGCCG	ATGCTGCCGT	GTTGGTGCTT	GCTTCCGGCA	CTGGTGAAAT	CGAAGCTGGT	ATTCGCAAGG	AA-----	[390]
057_EF1a_Jakoba_libera	GATTTTCATCA	AGAACATGAT	CACCGGTACC	TCTCAGGGTG	ATGCCGCCAT	CCTCATCGTC	GCCTCCGGCA	CTGGTGAGTT	CGAGGCCGGT	ATCTCCAAGG	AG-----	[390]
058_EF1a_Stachyamoeba_lipophora	GATTTTCATCA	AGAACATGAT	CACCGGTACC	TCTCAGGGTG	ATGTTGCCAT	TCTGATGATT	GATTCACAGA	CTGGTG6CTT	CGAGGCTGGT	TTCCGCCGTG	AC-----	[390]
059_EF1a_Naegleria_gruberi	GATTTTCATCA	AGAACATGAT	TACTGGTACC	TCCCAAGGGTG	ACGTTGCCAT	CTTGGTTGTT	GACTCCACTA	ACGGTG6TTT	CGAAGCTGGT	TTCTCTAAGG	AC-----	[324]
060_EF1a_Naegleria_andersoni	GATTTTCATCA	AGAACATGAT	TACTGGTACC	TCTCAAGCTG	ACGTTGCTAT	TTTGGTGATT	GACTCAACTA	ACGGTG6TTT	CGAAGCTGGT	TTCTCCAAGG	AC-----	[339]
061_EF1a_Acrasis_rosea	GATTTTCATCA	AGAACATGAT	CACCGGTACC	TACAGAAGCCG	ATGCCGCCGT	TTTGGTTATT	GATTTCAACA	CTGGTG6CTT	CGAAGCCGGT	ATCTCCAAGG	AC-----	[339]
062_EF1a_Sawyeria_marylandensis	GATTTTATTA	AAAATATGAT	CACCTGGTACT	TCACAAGCTG	ATGACGACAT	GTTAGTTATT	GATTACAGAA	AGGAGGTTT	TGAAGCAGGT	ATTGCAAGG	AA-----	[390]
063_EF1a_Bodo_saltans	GACTTCATCA	AGAACATGAT	TACCGGAACC	TCCCAAGGGTG	ATGCCGCCGT	CCTCGTGATT	GCCTCCGGCA	CTGGCGAGTT	CGAGGCTGGC	TTCTCCAAGG	AC-----	[336]
064_EF1a_Rhynchobodo_sp.	GACTTCATCA	AGAACATGAT	CACCGGCACC	TCCCAAGGGCG	ACGCCGCCAT	TCTAGTCACT	GCCTCCGGCG	AAGGTGAGTT	CGAGGCTGGT	ATCTCCAAGG	AA-----	[324]
065_EF1a_Rhynchomonas_nasuta	GACTTCATCA	AGAACATGAT	CACCGGCACC	TCTCAGGGCG	ACGCCGCCGT	GCTCATGATC	GCCTCCGGCG	AGGGCGAGTT	CGAGGCCGGC	TACGCGAAGG	AG-----	[387]
066_EF1a_Dimastigella_trypaniformis	GACTTCATCA	AGAACATGAT	CACCGGCACC	TGCGAGGGCG	ACGCCGCCAT	TCTGATCATC	GCCTCCGGCG	AGGGTGAGTT	CGAGGCTGGC	ATCTCCAAGG	AC-----	[345]

	340	350	360	370	380	390	400	410	420	430	440	
067_EF1a_Perkinsiella_amoebae	GACTTTCATCA	AGAACATGAT	CACGGGACG	TCCCAGGCGG	ACGCTGCCAT	CCTCGTCATT	GCCTCCGGCA	CCGGCGAGTT	CGAGGCCGGT	ATCAGCGGAG	AC-----	[315]
068_EF1a_Leishmania_major	GACTTTCATCA	AGAACATGAT	CACGGGACG	TCGCAAGCGG	ACGCGGCCAT	TCTGATGATC	GACTCGACGC	ATGGCGGCTT	CGAGGCTGGC	ATCTCGAAGG	AC-----	[390]
069_EF1a_Leishmania_infantum	GACTTTCATCA	AGAACATGAT	CACGGGACG	TCGCAAGCGG	ACGCGGCCAT	CCTGATGATC	GACTCGACGC	ATGGTGCGTT	CGAGGCTGGC	ATCTCGAAGG	AC-----	[390]
070_EF1a_Leishmania_braziliensis	GACTTTCATCA	AGAACATGAT	CACGGGACG	TCGCAAGCGG	ACGCGGCCAT	TCTGATGATC	GACTCGACGC	AGGGCGGCTT	CGAGGCGGGC	ATCTCGAAGG	AC-----	[390]
071_EF1a_Leishmania_tarentolae	GACTTTCATCA	AGAACATGAT	CACGGGACG	TCGCAAGCGG	ACGCGGCCAT	CCTGATGATC	GACTCGACGC	ACGGTGCGTT	CGAGGCTGGC	ATCTCGAAGG	AC-----	[390]
072_EF1a "L. tarentolae" FJ807257	GACTTTCATCA	AGAACATGAT	CACTGGTACC	TCCCAAGCGC	ACGCTGCGAT	CCTGATGGTT	GACTCGACGC	AGGGTGGGTT	TGAGGCTGGC	ATCTCGAAGG	AC-----	[315]
073_EF1a_Leptomonas_costaricensis	GACTTTCATCA	AGAACATGAT	CACGGGACG	TCGCAAGCGG	ATCGCGCCAT	TCTGATGATC	GACTCCACCC	AGGGTGCGTT	CGAGGCTGGC	ATCTCGAAGG	AC-----	[138]
074_EF1a_Leptomonas_bifurcata	GACTTTCATCA	AGAACATGAT	CACGGGACG	TCCCAGGCGG	ATGCGCGTAT	TCTGATGATC	GACTCGACCC	AGGGTGCGTT	CGAGGCTGGT	ATCTCGAAGG	AC-----	[150]
075_EF1a_Leptomonas_podlipaevi	GACTTTCATCA	AGAACATGAT	CACCGGACCC	TCCCAGGCGG	ACGCGCGCAT	TCTGATGATC	GACTCGACGC	AGGGTGCGTT	CGAGGCGGGT	ATCTCGAAGG	AT-----	[138]
076_EF1a_Herpetomonas_muscarum	GACTTTCATCA	AGAACATGAT	TACGGGACCC	TATCGGACCG	ATGCGGCCAT	CCTCATGATC	GACTCCACCC	AGGGTGCGTT	CGAGGCCGGC	ATTTCCAAGG	AC-----	[303]
077_EF1a_Herpetomonas_pessoai	GACTTTCATCA	AGAACATGAT	TACCGGTACC	TCCCAGGCGC	ATGCGCGTAT	TCTCATGATC	GACTCCACCC	AGGGTGCGTT	CGAGGCCGGT	ATTTCCAAGG	AT-----	[171]
078_EF1a_Blastocirithidia_culicis	GACTTTCATCA	AGAACATGAT	TACGGGACCC	TATCGGACCG	ATGCGCGGT	TCTGATGATC	GACTCCACCC	AGGGTGCGTT	TGAGGCTGGT	ATTTCCAAGG	AC-----	[303]
079_EF1a_Trypanosoma_brucei	GACTTTCATCA	AGAACATGAT	CACGGGACCA	TCGCAAGCGC	ACGCAAGCCAT	CCTCATCATT	GCCTCTGGCG	AGGGTGAGTT	CGAGGCTGGT	ATTTCCAAGG	AT-----	[390]
080_EF1a T. brucei equiperdum	GACTTTCATCA	AGAACATGAT	CACGGGACCA	TCGCAAGCGC	ACGCAAGCCAT	CCTCATCATT	GCCTCTGGCG	AGGGTGAGTT	CGAGGCTGGT	ATTTCCAAGG	AT-----	[312]
081_EF1a_Trypanosoma_aviium	GATTTCATCA	AGAACATGAT	TACGGGAACG	TCTCAGGCGG	ATCGGCCCAT	CCTCATCATT	GCCTCAGCGC	AGGGTGAGTT	CGAGGCTGGT	ATTTCCAARG	AC-----	[336]
082_EF1a_Trypanosoma_cruzi	GACTTTCATCA	AGAACATGAT	CACGGGACG	TCTCAGGCGG	ACGCGCGCGT	CCTTGTCATT	CGCTCATGCG	AGGGTGAGTT	TGAGGCTGGC	ATTTCCAAGG	AC-----	[390]
083_EF1a_Trypanosoma_rangeli	GACTTTCATCA	AGAACATGAT	CACTGGCACG	TCCCAGGCAG	ATGCGGCCAT	TCTGCTCATT	CGCTCTGGCG	AGGGCGAGTT	TGAGGCTGGT	ATTTCCAAGG	AC-----	[390]
084_EF1a_Ploetia_costatum	GACTTTCATCA	AGAACATGAT	CACCGGTACT	TCCCAGGCGG	ATGCGCGTGT	TCTTGTCATT	GACAGCACAA	CTGGTGCGTT	CGGAGCCGGT	ATTTCCAAGG	AT-----	[387]
085_EF1a_Ploetia_oblonga_cDNA	GACTTTCATCA	AGAACATGAT	CACCGGTACT	TCCCAGGCGG	ATGCGCGTGT	TCTTGTCATT	GACAGCACAA	CTGGTGCGTT	CGGAGCCGGT	ATTTCCAAGG	AT-----	[387]
086_EF1a_Ploetia_oblonga_gDNA	GACTTTCATCA	AGAACATGAT	CACGGGAACC	TCTCAGGCGC	ACGCTGCCGT	GTTGGTGATC	GATTGCACCA	CCGGTGCGTT	CGAGGCTGGT	ATTTCCAAGG	AT-----	[387]
087_EF1a_Entosiphon_sp.	GACTTTCATCA	AGAACATGAT	CACTGGTACC	TCCCAGGCGG	ATGCGCGGT	CCTTGTCATT	GACTCAACCA	CTGGTGCGTT	CGAGGCCGGT	ATCAGCAAGG	AT-----	[387]
088_EF1a_Entosiphon_sulcatum	GACTTTCATCA	AGAACATGAT	CACTGGTACC	TCCCAGGCTC	ACGCAAGCGT	CCTTGTCATT	GACTCAACCA	CTGGTGCGTT	CGAGGCCGGC	ATCAGCAAGG	AT-----	[315]
089_EF1a Peranema tricophorum	GACTTTCATCA	AGAACATGAT	TACTGGAACT	TCCCAAGCTG	ATGCTGCCGT	TTTGGTGATT	GACTCCACCA	CAGGAGGTTT	CGAAGCCGGT	ATCTCAAAGG	AT-----	[387]
090_EF1a P. tricophorum_FJ807243	GACTTTCATCA	AGAACATGAT	TACGGGAACG	TCCAAAGCTG	ATGCGCGTGT	CCTTGTCATT	GACTCCACCA	CAGGAGGTTT	CGAAGCTGGT	ATTTCAAAGG	AT-----	[381]
091_EF1a_Colacium_elongatum	GACTTTCATCA	AGAACATGAT	TACGGGAACC	TCCAGGCGG	ATGCTGCCGT	GCTGGTGATT	GACTCCACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCAAAGG	AT-----	[387]
092_EF1a_Colacium_mucronatum	GACTTTCATCA	AGAACATGAT	CACCGGGACT	TCCCAGGCTG	ATGCTGCCGT	GTTGGTGATT	GACTCCACCA	CTGGTGCGTT	TGAGGCTGGT	ATTTCAAAGG	AC-----	[333]
093_EF1a_Colacium_sp.	GACTTTCATCA	AGAACATGAT	TACCGGGACC	TCCCAGGCGG	ATGCTGCCGT	GCTGGTGATT	GACTCCACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCAAAGG	AT-----	[387]
094_EF1a_Colacium_vesiculosum	GACTTTCATCA	AGAACATGAT	CACGGGACT	TCCCAGGCTC	ATGCTGCCGT	GTTGGTGATT	GACTCCACCA	CTGGTGCGTT	TGAGGCTGGT	ATTTCAAAGG	AC-----	[345]
095_EF1a_Cryptoglena_pigra	GACTTTCATCA	AGAACATGAT	CACTGGGACA	TCCCAGGCTC	ATGCTGCCGT	CCTGGTGATT	GACTCCACCA	CTGGTGCGTT	CGAGGCGGGT	ATTTCAAAGG	AT-----	[387]
096_EF1a_Cryptoglena_skujai	GACTTTCATCA	AGAACATGAT	CACTGGGACA	TCCCAGGCTG	ATGCTGCCGT	CCTGGTGATT	GACTCCACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCAAAGG	AT-----	[387]
097_EF1a_Discoplastis_adunca	GACTTTCATCA	AGAACATGAT	CACTGGAACG	TCTCAGGCTC	ACGCTGCTGT	CCTGGTGATT	GACTCCACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCAAAGG	AT-----	[387]
098_EF1a_Euglena_clara	GACTTTCATCA	AGAACATGAT	CACTGGACG	TCTCAGGCGG	ATGCTGCTGT	GCTGGTGATT	GACTCCACCA	CCGGTGCGTT	CGAGGCTGGT	ATTTCAAAGG	AT-----	[267]
099_EF1a_Euglena_deses	GATTTCATCA	AGAACATGAT	CACTGGACCA	TCCCAGGCGG	ATGCTGCAGT	CCTGGTGATT	GACTCCACCA	CCGAGGCGCT	CGAGGCTGGC	ATTTCAAAGG	AT-----	[387]
100_EF1a_Euglena_exilis	GACTTTCATCA	AGAACATGAT	TACTGGAACC	TCCCAAGCTG	ATGCAAGTGT	CCTGGTGATT	GACTCCACCA	CTGGTGCGTT	CGAAGCTGGT	ATTTCAAAGG	AT-----	[267]
101_EF1a_Euglena_geniculata	GACTTTCATCA	AGAACATGAT	CACCGGTACT	TCCCAGGCGG	ATGCGCTGT	TCTTGTCATT	GACAGCACAA	CTGGTGCGTT	CGAAGCTGGT	ATTTCAAAGG	AT-----	[387]
102_EF1a_Euglena_gracilis	GATTTCATCA	AGAACATGAT	CACTGGTACC	TCCCAGGCTC	ACGCTGCAGT	GCTGGTGATT	GATTCCACCA	CTGGTGCGTT	CGAAGCTGGT	ATTTCAAAGG	AT-----	[387]
103_EF1a E. gracilis_X16890	GACTTTCATCA	AGAACATGAT	TACTGGACCC	TCCAGGCGTG	ATGCTGCAGT	GCTGGTGATC	GACTCCACCA	CAGGCTGGTT	GAAAGCTGGT	ATTTCAAAGG	AT-----	[390]
104_EF1a E. gracilis_FJ807238	GATTTCATCA	AGAACATGAT	CACTGGTACC	TCCCAGGCTC	ACGCTGCAGT	GCTGGTGATT	GATTCCACCA	CTGGTGCGTT	CGAAGCTGGT	ATTTCAAAGG	AT-----	[390]
105_EF1a_Euglena_granulata	GACTTTCATCA	AGAACATGAT	TACTGGACCA	TCGCAAGCTG	ATGCTGCTGT	CCTGGTTATC	GACAGCACAA	CTGGTGCGTT	TGAGGCTGGT	ATCTCGAAGG	AT-----	[387]
106_EF1a_Euglena_hiemalis	GATTTCATCA	AGAACATGAT	TACCGGACG	TCTCAGGCGG	ATGCTGCCGT	TTTGGTGATC	GACTCCACCA	CCGGTGCGTT	CGAGGCTGGT	ATCTCGAAGG	AC-----	[387]
107_EF1a_Euglena_longa												[0]
108_EF1a_Euglena_stellata	GATTTCATCA	AGAACATGAT	CACTGGCACT	TCTCAGGCTC	ATGCTGCCGT	TTTGGTTATT	GATTCTACCA	CTGGTGCGTT	GAAAGCTGGA	ATTTCAAAGG	AT-----	[267]
109_EF1a_Euglena_viridis	GATTTCATCA	AGAACATGAT	TACTGGCACT	TCTCAGGCGG	ATGCGWGTCT	CTTGGTGATT	GATTCCACCA	CCGGTGCGTT	TGAAAGCCGGT	ATTTCTAAGG	AT-----	[345]
110_EF1a_Eutreptia_pertyi	GACTTTCATCA	AGAACATGAT	TACTGGCAC	TCGCAAGCTG	ATGCTGCTGT	CCTGGTTATC	GACAGCACCA	CTGGTGCGTT	TGAGGCTGGT	ATTTCTAAGG	AT-----	[387]
111_EF1a_Eutreptiella_pomquetensis												[0]
112_EF1a_Lepocinlis_acus_var._major	GACTTTCATCA	AGAACATGAT	CACTGGCAGC	TCTCAGGCAG	ATGCTGCCGT	TCTCGTGATC	GACTCCACCA	CTGGTGCGTT	TGAGGCTGGT	ATTTCCAAGG	AT-----	[387]
113_EF1a_Lepocinlis_ovum	GACTTTCATCA	AGAATATGAT	CACTGGCACG	TCTCAGGCAG	ACGCGGCCGT	TTTGGTGATC	GACTCCACCA	CTGGTGCGTT	TGAAAGCTGGT	ATTTCTAAGG	AT-----	[267]
114_EF1a_Lepocinlis_spyrogyroides_clone_38												[0]
115_EF1a L. spyrogyroides_clones_1_&_4	GACTTTCATCA	AGAACATGAT	TACTGGTACC	TCTCAGGCTC	ACGCTGCCGT	GCTGGTTATC	GACTCCACGA	CTGGTGCGTT	CGAGGCTGGT	ATTTCTAAGG	AC-----	[387]
116_EF1a_Lepocinlis_steinii	GACTTTCATCA	AGAATATGAT	CACTGGCACG	TCTCAGGCAG	ACGCAAGCGT	CTTGGTGATC	GACTCCACCA	CTGGTGCGTT	TGAGGCTGGT	ATTTCCAAGG	AT-----	[267]
117_EF1a_Lepocinlis_tripteris	GACTTTCATCA	AGAACATGAT	CACTGGCAC	TCTCAGGCTC	ATGCTGCTAT	CCTGGTGATC	GATTCAACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCCAAGG	AT-----	[387]
118_EF1a_Monomorphina_aenigmatica	GACTTTCATCA	AGAACATGAT	CACAGGCACG	TCTCAGGCGG	ATGCTGCCGT	GCTGGTGATT	GACTCCACCA	CTGGTGCGTT	TGAGGCTGGT	ATTTCCAAGG	AC-----	[387]
119_EF1a_Monomorphina_pyrum	GACTTTCATCA	AGAACATGAT	TACTGGTACC	TCTCAGGCTG	ACGCTGCCGT	GCTGGTTATC	GACTCCACGA	CTGGTGCGTT	CGAGGCTGGT	ATTTCTAAGG	AC-----	[387]
120_EF1a_Monomorphina_sp.	GACTTTATCA	AGAACATGAT	CACTGGCACT	TCCCAGGCTC	ACGCTGCAGT	GCTGGTGATT	GACTCCACCA	CTGGTGCGTT	CGAAGCTGGT	ATTTCCAAGG	AT-----	[387]
121_EF1a_Phacus_orbicularis	GACTTTCATCA	AGAACATGAT	TACGGGCACG	TCTCAGGCAG	ACGCTGCTGT	GCTGGTGATT	GACTCCACCA	CTGGTGCGTT	TGAGGCTGGT	ATTTCCAAGG	AT-----	[345]
122_EF1a_Phacus_similis	GATTTCATCA	AGAACATGAT	CACTGGCACG	TCTCAGGCAG	ACGCGCCGCT	TCTGGTGATC	GACTCCACCA	CTGGTGCGTT	TGAGGCAAGT	ATTTCCAAGG	AC-----	[387]
123_EF1a_Strombomonas_acuminata	GACTTTCATCA	AGAATATGAT	TACGGGACCA	TCCCAGGCTC	ATGCAAGCGT	CTTGGTCATT	GACTCCACCA	CTGGTGCGTT	TGAAGCTGGT	ATTTCCAAGG	AT-----	[390]
124_EF1a_Strombomonas_borsteniensis	GACTTTCATCA	AGAACATGAT	CACTGGGACT	TCCCAGGCTG	ATGCTGCTGT	GCTGGTTATC	GACTCTACCA	CTGGTGCGCT	TGAGGCTGGT	ATTTCCAAGG	AT-----	[267]
125_EF1a_Strombomonas_napiformis	GATTTCATCA	AGAACATGAT	CACTGGGACT	TCTCAGGCTC	ACGCTGCTGT	GCTGGTTATC	GACTCCACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCCAAGG	AT-----	[387]
126_EF1a_Strombomonas_ovalis	GATTTCATCA	AGAACATGAT	CACTGGGACT	TCTCAGGCTC	ATGCTGCTGT	GCTGGTTATC	GACTCCACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCCAAGG	AT-----	[387]
127_EF1a_Strombomonas_triquetra	GATTTCATCA	AGAACATGAT	TACAGGGACC	TCCCAGGCGG	ATGCTGCCGT	GCTGGTGATT	GATTCTACTA	CTGGTGCGTT	CGAGGCTGGT	ATTTCCAAGG	AC-----	[387]
128_EF1a_Trachelomonas_conradi	GACTTTCATCA	AGAACATGAT	TACTGGTACC	TCCCAGGCTG	ACGCGCCGCT	GCTGGTGATT	GATTCCACCA	CTGGTGCGTT	CGAGGCAAGT	ATTTCCAAGG	AT-----	[387]
129_EF1a_Trachelomonas_echinata	GATTTCATCA	AGAACATGAT	TACTGGCAC	TCTCAGGCTC	ATGCTGCCGT	TTTGGTGATC	GATTCCACCA	CTGGTGCGTT	TGAAGCAAGT	ATTTCCAAGG	AT-----	[387]
130_EF1a_Trachelomonas_levfveii	GACTTTATCA	AGAATATGAT	CACTGGGAAG	TCTCAGGCTG	ATGCGCAGGT	GTTGGTGATC	GACTCCACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCCAAGG	AT-----	[387]
131_EF1a_Trachelomonas_plancitonica	GACTTTATCA	AGAATATGAT	CACTGGGACG	TCTCAGGCTC	ACGCTGCAGT	GTTGGTGATC	GACTCTACCA	CTGGTGCGTT	CGAGGCTGGT	ATTTCCAAGG	AT-----	[267]
132_EF1a_Trachelomonas_similis	GACTTTCATCA	AGAACATGAT	CACAGGGACC	TCCAGGCGTG	ACGCTGCCGT	TCTTGTCATC	GATTCCACCA	CTGGTGCGTT	TGAAGCTGGT	ATTTCCAAGG	AT-----	[267]

	340	350	360	370	380	390	400	410	420	430	440	
133_EFL1a_Trachelomonas_volvocina	GATTTTCATCA	AGAACATGAT	CACAGGGAGT	TCACAGGGCT	ATGCTGCCGT	GTTGGTCAAT	GATTCAACCA	CTGGTGGTTT	TGAGGCTGGT	ATTTTCAAGG	AT-----	[387]
134_EFLenvironmental_sample_EU810232	GATTTTCATCA	AGAACATGAT	CACCGGGATC	AGCCAGGCCG	ATGTGGCCAT	GGTTCTGGTG	CCCGCCGACG	GAAACTTCAC	CACCAAGCATT	GCC---AAGG	GCAACCCACAA	[395]
135_EFLCreolimax_fragrantissima	GATTTTCATTA	AGAACATGAT	CACTGGTGCT	TCCCAGGCTG	ATGTCGCGCT	TCTTATGGTT	CCTTCTGACG	GTAACCTTCAG	TGCCGCCATT	GCC---AAGG	GTAACCCACAA	[323]
136_EFLSphaeroforma_arctica	GATTTTCATCA	AGAACATGAT	CACTGGTGCT	TCTCAGGCTG	ATGTGTCCT	ACTTATGGTG	CCATCTGACG	GTAACCTTCAG	TGCTGCCATT	CCC---AAGG	GTAATCACAA	[395]
137_EFLMonosiga_brevicollis	GACTTCATCA	AGAACATGAT	CACCTGGTCT	TCCCAGGCCG	ATGTTGCCCT	GCTCATGGTT	CCCTGCGATG	GTAACCTTCAC	TGCTGCCATT	GCC---AAGG	GCAACCCACAA	[395]
138_EFLXenopus_laevis	GATTTTCATCA	AGAACATGAT	CACGGGTGCC	TCCCAGGCCG	ATGTGCCCCT	TCTGATGATC	CCCGCCGACG	GTAACCTTCAC	CGCCGCCATC	GCC---AAGG	GCAACCCACAA	[395]
139_EFLSpizellomyces_punctatus	GATTTTCATCA	AGAACATGAT	CTCTGGTGCC	GCCCAGGCCG	ATGTGCGTCT	TCTCATGGTT	CCGCGTGACG	GTAACCTTCAC	CACTGCCATC	CAG---AAGG	GTGACCACAA	[395]
140_EFLGaertneriomyces_semiglobiferous	-----A	AGAACATGAT	CTCTGGTGCT	GCCCAGGCCG	ATGTGCGTCT	TCTCATGGTT	CCGCGTGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GTGACCACAA	[98]
141_EFLAllomyces_macrogynus	GATTTTCATCA	AGAACATGAT	CTCGGGTGCC	GCCCAGGCCG	ATGTGCGTCT	CCTCATGGTC	CCGCGCGACG	GTAACCTTCAC	GACGGCCATC	CAG---AAGG	GTGACCACAA	[395]
142_EFLBlastocladiella_emersonii	GACTTCATCA	AGAACATGAT	CTCGGGTGCC	GCCCAGGCCG	ATGTGCGCCT	CCTCATGGTC	CCGCGCGACG	GTAACCTTCAC	CACGGCTATC	CAG---AAGG	GTGACCACAA	[395]
143_EFLCatenophylctis_sp.	-----A	AGAACATGAT	TTCCGGGCGT	GCCCAGGCCG	ATGTGGCCCT	CCTCATGGTG	CCGCGTGACG	GTAACCTTCAC	CACGGCCATC	CAA---AAGG	GTGACCACAA	[98]
144_EFLCoelomomyces_stegomyiae	-----A	AGAATATGAT	TTCTGGTGCT	GCCCAAGCTG	ATGTTGCCCT	GTTGATGGTG	CTGCTGATG	GTAACCTTCAC	CGTGCCATT	CAA---AAGG	GTGACCACAA	[98]
145_EFLEntomophthora_muscae	-----A	AGAACATGAT	TTCCGGAGCC	GCCCAGGCTG	ATGTGCCCCT	TCTCATGGTC	CCCGCCGACG	GTAACCTTCAC	CACCGCCATC	CAA---AAGG	GTGACCACAA	[98]
146_EFLBasidiobolus_ranarum	-----A	AGAACATGAT	TTCCGGAGCC	GCTCAAGCTG	ATGTGYCYCT	TCTCATGGTY	CCCGCCGATG	GTAATTTTAC	CACCGCYATC	CAA---AAGG	GTGACCACAA	[98]
147_EFLConidiobolus_coronatus	-----A	AGAACATGAT	TTCCGGTGCC	GCTCAAGCTG	ATGTGCTTTT	ATTAATGGTT	CCGCGTGACG	GTAACCTTCAC	CACCTGCCATC	CAA---AAGG	GTGACCACAA	[98]
148_EFLTetraselmis_tetrathele	GATTTTCATCA	AGAACATGAT	TTCTGGTGCC	GCCCAGGCCG	ATGTGTGCCCT	GCTGATGGTG	CCGCGCGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GCAACCCACAA	[344]
149_EFLMesostigma_viride	GATTTTCATCA	AGAACATGAT	TTCCGGTGCT	GCCCAGGCCG	ATCTGGCCCT	GCTGATGGTG	CCGCGTGATG	GTAACCTTCAC	CACCTCTATC	CAG---AAGG	GTGACCACAA	[398]
150_EFLOstreococcus_tauri	GATTTTCATTA	AGAACATGAT	TTCCGGTGCC	GCCCAGGCCG	ACGTGCCCCT	CCTCATGGTC	CCGGCCGACG	GTAACCTTCAC	CGTCGCCATC	CAA---AAGG	GTAACCCACAA	[395]
151_EFLNephrolepis_olivacea	GATTTTCATCA	AGAACATGAT	CTCTGGTGCT	GCCCAGGCCG	ACGTGGCTCT	GCTCATGGTG	CCCGCCGATG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GTGACCACAA	[371]
152_EFLMicromonas_pusilla	GATTTTCATCA	AGAACATGAT	CTCCGGGCGC	GCGCAGGCTG	ACGTGCCCCT	CCTCATGGTC	CCCGCCGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GCAACCCACAA	[344]
153_EFLParachlorella_kessleri	GATTTTCATCA	AGAACATGAT	CTCCGGTGCC	GCCCAGGCCG	ACGTGTGCCCT	GCTCATGGTG	CCCGCCGACG	GTAACCTTCAC	GACCGCCATC	CAG---AAGG	GTGACCACAA	[353]
154_EFLHelicosporidium_sp.	GATTTTCATCA	AGAACATGAT	TTCCGGTGCC	GCCCAGGCCG	ACGTGTGCTT	GCTGATGGTG	CCCGCCGACG	GTAACCTTCAC	TACCGCCATC	CAG---CGCG	GCACCCACAA	[395]
155_EFLGonium_pectorale	GACTTCATCA	AGAACATGAT	TTCCGGTGCT	GCCCAGGCCG	ATGTGTGCCCT	GCTGATGGTG	CCCGCCGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GTGACCACAA	[395]
156_EFLChlamydomonas_reinhardtii	GACTTCATCA	AGAACATGAT	TTCCGGTGCC	GCCCAGGCCG	ATGTGTGCCCT	GCTGATGGTG	CCCGCTGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GTGACCACAA	[395]
157_EFLScenedesmus_obliquus	GACTTCATCA	AGAACATGAT	CTCTGGTGCT	GCCCAGGCTG	ACGTGTGCCCT	GCTGATGGTG	CCCGCCGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GTGACCACAA	[395]
158_EFL"Oryza_sativa"	GATTTTCATCA	AGAACATGAT	CTCTGGTGCT	GCTCAGGCTG	ATGTGTGCCCT	GCTGATGGTG	CCCGCTGATG	GTAACCTTCAC	CACCTGCCATT	CAG---AAGG	GTGACCACAA	[395]
159_EFLChlorococcum_sp.	GATTTTCATCA	AGAACATGAT	CTCAGGCGCC	GCCCAGGCTG	ATGTCTGTCT	TCTCATGGTT	CCCGCTGACG	GTAACCTTCAC	CACCTGCCATC	CAG---AAGG	GTGACCATAA	[344]
160_EFLOchlochaete_hystrix	GATTTTCATCA	AGAACATGAT	CTCCGGTGCC	GCCCAGGCCG	ATGTGTGCCCT	GCTCATGGTG	CCCGCCGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GTGATTCACAA	[353]
161_EFLUlvula_intestinalis	GATTTTCATCA	AGAACATGAT	TTCTGGGCGT	GCTCAGGCTG	ACGTCTGTCT	GCTGATGGTG	CCCGCTGATG	GTAACCTTCAC	CACCTGCTATC	CAG---AAGG	GTGACCACAA	[344]
162_EFLUrospora_sp.	GACTTCATCA	AGAACATGAT	CTCGGGGCGC	ACGTCTGCGCT	CCTCATGGTG	CTCATGGTG	CCCGCCGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GCACCCACAA	[197]
163_EFLGymnochlorella_stellata_ppc	GATTTTCATTA	AGAATATGAT	TTCCGGTGCT	GCCCAGGCAG	ATGTAGCTCT	CTTATGCTT	CCCGCTGATG	GTAACCTTCAC	GACTGCGATT	CAG---AAGG	GAAACCATAA	[431]
164_EFLLotharella_amoebiformis_ppc	GATTTTCATCA	AAAACATGAT	TTCCGGTGCC	GCCCAGGCCG	ACGTAGCCCT	ATTAATGCTC	CCCGCCGACG	GTAACGCATG	GACGGCTATC	CAA---AAGG	GAATTCACAA	[176]
165_EFLChlorarachnion_reptans_ppc	GACTTCATCA	AAAACATGAT	TTCTGGTGCC	GCCCAGGCCG	ACGTGTGCCCT	GCTGATGCTT	CCTGCTGACG	GTAACGTTG	GACTGCCATC	CAG---AAGG	GAGATCACAA	[203]
166_EFLChondrus_crispus	GATTTTCATCA	AGAACATGAT	TACTGGTGCC	TCCCAGGCCG	ACGTGGAGAT	TCTCATGGTT	CCCGCCGATG	GAAACTTTAT	TACTTCTATT	GCC---AAGG	GCATTCACAA	[395]
167_EFLGracilaria_changii	GATTTTCATCA	AGAACATGAT	TAAGTGGTCC	TACAAGGCTG	ATGTGTGCTC	CCTCATGGTT	CCCGCCGATG	GTAATTTTAC	CACCTTCCATC	GCC---AAGG	GTGACCATAA	[395]
168_EFLCryptomonas_ovata	GATTTTCATCA	AGAACATGAT	CTCTGGGCGT	GCCCAGGCCG	ACGTGCGTCT	TCTCATGGTT	CCCGCTGATG	GCAACTTTG	GACCGCCATC	CAA---AAGG	GTGATGCCAA	[350]
169_EFLGuillardia_theta	GATTTTCATCA	AGAACATGAT	CTCCGGATCC	GCTCAGGCTG	ACGTGCGTCT	TCTCATGGTT	CCCGCTGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GAGATGCCAA	[392]
170_EFLRhodomonas_salina	GATTTTCATCA	AGAACATGAT	CTCTGGGCGC	GCCCAGGCTG	ACGTGGCCGT	GCTCATGGTT	CCCGCTGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GTGATGCCAA	[392]
171_EFLPavlova_lutheri	GACTTTATCA	AGAACATGAT	CTCGGGGCGC	GCGCAGGCCG	ACGTGCGCCT	GCTGATGGTG	CCGCGGACG	GCAATTTTAC	CACGGCGATC	CAG---AAGG	GCAACCCACAA	[392]
172_EFLChrysochromulina_sp.	GACTTCATCA	AGAACATGAT	TTCCGGTGCC	GCCCAGGCCG	ATGTGCGCCT	CCTCATGGTC	CCGCGCGACG	GTAACCTTCAC	CACCGCCATC	CAG---AAGG	GCAACCCACAA	[212]
173_EFLIsochrysis_galbana	GACTTCATCA	AGAACATGAT	CTCTGGGCGA	GCCCAGGCCG	ATGTGCCCCT	CCTCATGGTC	CCGCGCGACG	GTAACCTTCAC	CACGGCAATC	CAG---AAGG	GCAACCCACAA	[395]
174_EFLEmiliana_huxleyi	GACTTCATCA	AGAACATGAT	CTCCGGGCGC	GCGCAGGCCG	ACGTGGCCCT	CCTCATGGTG	CCCGCCGACG	GCAACTTCAC	CACCGCCATC	CAG---AAGG	GCAACCCACAA	[392]
175_EFLAchnanthes_kuwaitensis	GATTACATCA	AGAACATGAT	CACCGGATCT	GCCACGCCG	ATGTGCGTCT	TCTTCTCGTT	CCCGCCGAGA	AGGGTGTTT	CGAGGCCGCT	ATCGCGAAGG	CTGATGCCAA	[332]
176_EFLAsterionella_glacialis	GATTACATCA	AGAACATGAT	TACCGGATCT	GCCACTGCCG	ATGTTGGTCT	TCTTCTTGTT	CCGCGTGAGA	AGGGTGTTT	CGAGGCTGCC	ATTGCCAAGG	CCGACGCCAA	[332]
177_EFLDetonula_confervacea	GATTACATCA	AGAACATGAT	TACCGGATCT	GCCACGCCG	ATGTTGGACT	TCTTCTTGTC	CCCGCCGAAA	AGGGTGTTT	CGAGGCTGCC	ATTGCCAAGG	CCGATCCCAA	[332]
178_EFLThalassiosira_pseudonana	GATTTTCATCA	AGAACATGAT	CACCTGGATCT	GCCACCGCTG	ATGTTGGTCT	TCTTCTTGTT	CTGCGCGAGA	AGGGAGGTTT	CGAGGCCGCT	ATCGCCAAGG	CCGATGCCAA	[395]
179_EFLPerkinsus_marinus	GATTTTCATCA	AGAATATGAT	TACTGGTGCA	TCCCAGGCAG	ATGTGCTCT	TCTCATGGTT	CTGCTGATG	GTAACCTTTG	CACCTGCTATC	GCC---AAGG	GCAATCACAA	[404]
180_EFLOxyrrhis_marina	GATTTTCATCA	AGAACATGAT	TTCCGGAGCT	TGCGAGGCCG	ACGTGGCCCT	GATCATGGTT	CCGGCCGATG	GTAACCTTCAT	CGTGTCTATC	CAG---AAGG	GCAACCCACAA	[401]
181_EFLKarlodinium_micrum	GACTTTATCA	AAAACATGAT	TACTGGTGCT	TGCGAAGCCG	ATGTGCGCTT	GATCATGGTT	CCAGCAGATG	GTAACCTTCAC	CACCTGCCATC	GCA---AAGG	GCAACCCACAA	[401]
182_EFLAmphidinium_carterae	GACTTCATCA	AGAACATGAT	CACAGGCCGCT	TCCCAGGCTG	ACGTGCGCCT	CATCATGGTT	CCAGCAGATG	GTAACCTTCAC	CACCTGCTATT	GCC---AAGG	GCAACCCACAA	[401]
183_EFLHeterocapsa_triquetra	GACTTCATCA	AGAACATGAT	CACCGGTGCC	TGCGAGGCTG	ACGTGGCCCT	CATCATGGTT	CCCGCCGATG	GTAACCTTCAC	CACCTGCCATT	GCC---AAGG	GCAACCATAA	[401]
184_EFLPlanoglabratella_operculis	GACTTTATCA	AGAACATGAT	CTCCGGTGCT	TCCCAGGCCG	ATGTGCAATT	GTTAATGGTC	CTGCGCAACA	AAGGAGGCTT	TGAGACCTCC	ATTGCCAAGG	GAAACCATAA	[233]
185_EFLReticulomyxa_filosa	GATTTTATTA	AGAACATGAT	CACGGGAGCT	TGCGAAGCTG	ATGTGCGTCT	GCTGATGGTG	CTGCGCAACA	AGGGTGATTT	TGAAACCTCA	ATTGCCAAGG	GAAACCCACAA	[323]
186_EFLLotharella_amoebiformis	GACTTCATCA	AAAACATGAT	TTCCGGTGCC	TGCGAGGCCG	ACGTAGCCGT	TCTCATGGTG	CCGCGCAACA	AGGGTGTTT	CGAGGTTGCC	ATTCAGAAAG	GTGATGGTAC	[440]
187_EFLGymnochlorella_stellata	GACTTCATCA	AGAACATGAT	TTCCGGAGCT	TCTCAAGCAG	ATGTTGCCCT	TCTTATGGTT	CCGCGCAAGA	AAGGTGGTTT	CGAAGCTGCT	ATTCAGAGGG	GAGATGGAACT	[356]
188_EFLChlorarachnion_reptans	GACTTCATCA	AAAACATGAT	CTCCGGTGCT	TCCCAGGCCG	ATGTGCCCCT	TCTTATGGTT	CTGCGCAAGA	AGGGTGTTT	TGAGGCTGCC	ATCCAGAAAG	GTGATGGTTT	[323]
189_EFLDiplonema_ambulator	-----A	-----A	-----A	-----A	-----A	-----A	-----A	-----A	-----A	-----A	-----A	[0]
190_EFLDiplonema_papillatum	GACTTCATCA	AGAACATGAT	CACCGGTGCC	TGCGAGGCTG	ACGTGCCCCT	GCTGATGATC	CCTGCCGACG	GTAACCTTCG	GTCCGCGATT	GCC---CGTG	GTAACCCACAA	[368]
191_EFLRhynchopus_euleeides	-----A	-----A	-----A	-----A	-----A	-----A	-----A	-----A	-----A	-----A	-----A	[8]
192_EFLPetalomonas_cantuscygni	GATTTTCATCA	AGAACATGAT	CACCGGTGCC	TCCCAGGCCG	ACGTGCGCAT	CCTCTCTGCT	CCGGCCGATG	GTAACCTTCAC	CACCGCCATC	GCC---AAGG	GTGACCACAA	[239]
193_EFLTrypanoplasma_borrelia	GAYTTYATYA	AGAACATGAT	MAGTGGTCTC	TCCCAGGCCG	ACGTGGCTTT	ACTGCTGTG	CCYGCTGACG	GYGGWAGCTT	CATTGCCATC	ATTGCCAAGG	GCACCCACAA	[404]
194_EFLNeobodo_salienis	GACTTCATCA	AGAACATGAT	CACCGGTGCC	TCCCAGGCCG	ACGTGCCCCT	GCTCTCTGTC	CCGGCCGACG	GTGGCTCGTT	CATCGCGTCC	ATCGCCAAGG	GCACCCACAA	[353]

	450	460	470	480	490	500	510	520	530	540	550	
001_aEF1a_Desulfurococcus_mobilis	-----	GGCC	AGACACGTGA	GCACGCTATA	CTAGCGAGAA	CCATGGGTAT	AAACCAAGTAT	ATCGTCGCGA	TAAACAAGAT	GGATGCCACT	GAGCCACCG-	[486]
002_aEF1a_Sulfolobus_solfataricus	-----	GGAC	AAACAAGAGA	GCATATAATT	CTAGCCAAGT	CCATGGGTAT	AGATCAATTG	ATAGTTGCAG	TAAACAAAT	GAGCTTAACA	GAGCCACCA-	[480]
003_aEF1a_Thermoplasma_acidophilum	-----	GaGc	AgAcCaGaGa	GcAtGcCtTc	CtTgCgAgGa	CtCtTgTgTg	AcCaCaGaTg	GtTgTgGcCa	TaAaCaAgAt	GgAtGcGaCa	AgCccCaCcG-	[468]
004_aEF1a_Halobacterium_sp.	-----	CCCC	AGACCCGAGA	GCACGTCTTC	CTCTCGCGAA	CGCTTGGCAT	CGACGAGCTC	ATCGTCGCTG	TCAACAAGAT	GGAC-----	GTGCTGAC-	[453]
005_aEF1a_Methanocaldococcus_jannaschii	-----	CCAC	AAACAAGAGA	GCACATGTTT	TTAGCAAGAA	CATTGGGTAT	TAAGCAAAAT	GCAGTTGCAA	TTAACAAGAT	GGAT-----	ACAGTTAAC-	[465]
006_EF1a_Acanthamoeba_culbertsoni	-----	GGTC	AGACCCGTGA	GCACGCCCTG	CTCGCTTTCA	CTCTCGGTGT	CAAGCAGATG	ATCGTTGTGT	GCAACAAGAT	GGAC-----	AACGTCAAC-	[450]
007_EF1a_Neoparamoeba_branchipila	-----	GGTC	AAACCCGTGA	GCACGCCCTG	CTCTCCTTCA	CCCTCGGTGT	CAAGCAAATG	ATCTGCGCCA	TCAACAAGAT	GGATGACAAG	TCTGTCAAC-	[435]
008_EF1a_Planoprotostelium_aurantium	-----	GGAC	AGACCAGAGA	GCACGCTCTG	CTTGCTTCA	CCCTTGGTGT	CAAGCAGCTC	ATCGTCTCTG	TCAACAAGAT	GGACGAGAAG	TCCACCAAC-	[432]
009_EF1a_Physarum_polycephalum	-----	GGAC	AGACCCGTGA	GCACGCTCTC	CTTGCTATA	CCTTGGGTGT	GAAGCAAATG	ATCGTCGCCA	TCAACAAGAT	GGATGAGAAG	TCCGTCAAC-	[432]
010_EF1a_Amoebidium_parasiticum	-----	GGCC	AGACCCGTGA	GCACGCTCTG	CTCGCTTCA	CCCTCGGTGT	CAAGCAGCTG	ATCGTTGCCA	TCAACAAGAT	GGAC-----	TCGATCAAG-	[447]
011_EF1a_Capsaspora_owczaraki	-----	GGCC	AGACCCGTGA	GCACGCCCTG	CTCGCTTCA	CCCTCGGTGT	CAAGCAGCTG	ATGTCGCCA	TCAACAAGAT	GGAC-----	TCGATCAAG-	[477]
012_EF1a_Monosiga_ovata	-----	GGCC	AAACCCGCGA	GCACGCCCTC	CTCGCTTCA	CCCTCGGTGT	CAAGCAGCTC	ATCGTCGGTG	TCAACAAGAT	GGACGCACT	GAGCCCCCG-	[456]
013_EF1a_Salpingoeca_amphoridium	-----	GGCC	AGACCCGTGA	GCACGCCCTG	CTCGCTTCA	CCCTGGGTGT	CAAGCAGCTC	ATGTCGCCG	TGAACAAGAT	TGACTCGACT	GAGCCCCCG-	[450]
014_EF1a_Ceanorhabditis_elegans	-----	GGAC	AAACTCGCGA	GCACGCTCTT	CTTGCCAGA	CCCTTGAAGT	CAAGCAACTC	ATCGTTGCCT	GCAACAAGAT	GGACTCCACC	GAGCCACCA-	[483]
015_EF1a_Drosophila_melanogaster	-----	GGCC	AGACCCGCGA	GCACGCCCTG	CTCGCTTCA	CCCTGGGTGT	GAAGCAGCTG	ATCGTTGGTG	TGAACAAGAT	GGACTCTCCG	GAGCCACCA-	[483]
016_EF1a_Ornithorhynchus_anatinus	-----	GGGC	AGACTCGTGA	GCATGCCCTG	TTGGCTTACA	CCTTGGGGTG	GAACAGACTC	ATTGTGGGGA	TCAACAAGAT	GGACTCCACA	GAGCCCCCG-	[483]
017_EF1a_Xenopus_laevis	-----	GGAC	AGACCCGGGA	GCATGCTCTC	TTGGCTTCA	CCCTAGGTGT	CAAGCAGCTT	ATAATTGGAG	TTAACAAGAT	GGACTCCACT	GAGCTCTCT-	[483]
018_EF1a_Saccharomyces_cerevisiae	-----	GGTC	AAACCAGAGA	ACACGCTTTG	TTGGCTTCA	CCTTGGGTGT	TAGACAATTG	ATTGTTGCTG	TCAACAAGAT	GGAC-----	TCCGTCAAA-	[477]
019_EF1a_Neurospora_crassa	-----	GGCC	AGACCCGTGA	GCACGCCCTG	CTCGCTTCA	CCCTCGGTGT	CAAGCAGCTC	ATTGTTGCCA	TCAACAAGAT	GGAC-----	ACCCGCAAG-	[480]
020_EF1a_Cryptococcus_neoformans	-----	GGTC	AGACCCGCGA	GCACGCCCTC	CTCGCTTCA	CCCTCGGTGT	CAGGCACTG	ATTGTTGCTT	GCAACAAGAT	GGAC-----	ACCTGCAAG-	[477]
021_EF1a_Schizophyllum_comune	-----	GGCC	AGACCCGCGA	GCACGCCCTC	CTTGCTTCA	CCCTCGGTGT	CGCTGAGCTC	ATCGTCGCCG	TCAACAAGAT	GGAC-----	ACGACCAAG-	[477]
022_EF1a_Monoblepharis_insignis	-----	GGTC	AGACCCGTGA	GCACGCTCTG	CTCTGCTTCA	CGCTCGGTGT	CGCTGAGCTC	ATTGTTGCCG	TGAACAAGAT	GGAC-----	AGCATGGC-	[363]
023_EF1a_Rhizophylyctis_harderi	-----	GGTC	AAACCCGTGA	ACACGCTCTT	CTTGCCAACA	CCCTCGGTGT	CAAGCAGATG	ATCATTTGCTG	TCAACAAGAT	GGAC-----	ACCAACAAG-	[363]
024_EF1a_Harpochytrium_sp.	-----	GGTC	AGACCCGTGA	GCACGCCCTT	CTCGCTTCA	CCCTCGGGTG	GAAGCAGATG	ATCGTCGGTG	TCAACAAGAT	GGAC-----	TCCGCCAAG-	[315]
025_EF1a_Absidia_glauca	-----	GGTC	AAACCCGTGA	ACACGCTTTG	TTGGCTTCA	CCCTCGGTGT	CCGTCAATTG	ATCGTCGCCA	TTAACAAGAT	GGAT-----	TCCACCAAG-	[477]
026_EF1a_Smittium_simulii	-----	GGTC	AAACCCGTGA	GCACGCTCTT	CTTGCTTCA	CCTTGGGTGT	CAGACAATT	ATTGTTGCTG	TCAACAAGAT	GGAT-----	TCTAACAAA-	[450]
027_EF1a_Chara_australis	-----	GGCC	AGACCCGTGA	GCATGCTCTG	TTGGCTTCA	CGCTGGGCGT	GAAGCAGATG	ATCTGCGCTT	GCAACAAGAT	GGATGCCACA	ACGCCCAAG-	[432]
028_EF1a_Physcomitrella_patens	-----	GGGC	AGACCCGTGA	GCAGCAGCTG	CTGGCGTTCA	CTTTGGGAGT	GAAGCAGATG	ATTGCTGCTT	GCAACAAGAT	GGACGCGACC	ACCCGCAAG-	[483]
029_EF1a_Acetabularia_acetabulum	-----	GGCC	AAACTCGCGA	ACACGGCTCT	CTAGCAATTCA	CACCTAGGAGT	AAAGCAGATG	ATTGTTGGGT	GCAACAAGAT	GGACGCCACC	GAGCTCTCCC-	[432]
030_EF1a_Chaetomorpha_coliformis	-----	GGTT	AGACGAGAGA	GCATGCTCTG	CTTGCTTTA	CCCTGGGTGT	CAAGCAGATG	ATCTGCTGCC	CTAATAAAAT	GGACGCTACT	GAGCCCCCA-	[282]
031_EF1a_Arabidopsis_thaliana	-----	GGTC	AGACCCGTGA	GCACGCTCTT	CTTGCTTCA	CCCTTGGGTGT	CAAGCAGATG	ATTGCTGCTT	GTAACAAGAT	GGATGCCACC	ACCCCAAA-	[483]
032_EF1a_Oryza_sativa	-----	GGCC	AGACCCGTGA	GCACGCTCTT	CTTGCTTCA	CCTTGGGAGT	GAAGCAGATG	ATCTGCTGCC	GCAACAAGAT	GGATGCCACT	ACTCCAAG-	[483]
033_EF1a_Cyanidioschyzon_merolae	-----	GGAC	AAACTCGTGA	ACACGCGCTG	CTCGCTTCA	CGCTCGGTGT	GAAGCAGATG	ATTGTTGGCG	TGAACAAGAT	GGACGATAAG	AACGTGAAC-	[483]
034_EF1a_Porphyrha_jezoensis	-----	GGGC	AGACCCGCGA	GCACGCGCTG	CTGGCTTCA	CCCTGGGCGT	CAAGCAGATG	ATTGTTGGCT	GCAACAAGAT	GGACGACAAG	AACGTCAAC-	[483]
035_EF1a_Plasmodium_falciparum	-----	GGTC	AAACAAAGGA	ACACGCTTTA	TTAGCAATTG	CCTTAGGTTG	TAAGCAAAAT	GTTGTCGGTG	TTAACAAGAT	GGAT-----	ACTGTCAAA-	[477]
036_EF1a_Cryptosporidium_parvum	-----	GGTC	AAACAGAGA	ACATGCTTTA	TTGGCTTCA	CTTTGGGTGT	CAGACAATTG	ATTGTCGGTA	TTAACAAGAT	GGAT-----	ACCTGTGAA-	[471]
037_EF1a_Tetrahymena_pyriformis	-----	GGTC	AAACAGAGA	ACACGCTCTG	TTGGCTTCA	CTTTGGGTGT	CAAGCAAATG	ATCGTCTGCC	TTAACAAGAT	GGACGAGAAG	ACCGTCAAC-	[486]
038_EF1a_Stentor_coeruleus	-----	GGCC	AGACCAAGGA	ACACGGTCTA	TTAGCTTCA	CCCTTGGGTGT	CAAGCAAATG	ATTGTTGGCG	TCAACAAGAT	GGACGACAAG	TCTTGCAAT-	[438]
039_EF1a_Phytophthora_parasitica	-----	GGCC	AGACCCGTGA	GCACGCTCTG	CTTGCTTCA	CCTTGGGTGT	GAAGCAGATG	ATCGTCGCCA	TCAACAAGAT	GGACGATG	TCTGTCTATG-	[483]
040_EF1a_Mallomonas_rasilis	-----	GGTC	AGACCAAGGA	GCATGCACTT	TTGGCTTTA	CCCTTGGGTGT	GAACCAAATG	ATTGTTGCTA	TTAACAAGAT	GGAT-----	GGCTGCAAC-	[444]
041_EF1a_Phaeodactylum_tricornutum	-----	GGAC	AGACCCGCGA	ACACGCGCTT	CTCGCTTCA	CTCTTGGAGT	GAAGCAGATG	ATTGTTGCCA	TGAACAAGAT	GGACGACAAG	ACTGTCAAA-	[483]
042_EF1a_Thalassiosira_pseudonana	-----	GGCC	AGACGAGAGA	ACATGCAATT	TTGGCTTCA	CACCTGGGAGT	ACGACAAGTT	ATTGTTGTTG	TGAACAAGAT	GGACGATAAG	TCGGTACAG-	[483]
043_EF1a_Malawimonas_jakobiformis	-----	GGCC	AGACCCGCGA	GCACGCCCTG	CTTGCTTCA	CTCTTGGGTGT	GAAGCAGATG	GTCGTTGCCG	TCAACAAGAT	GGACGACAAG	ACCGTCAAC-	[414]
044_EF1a_Malawimonas_californiana	-----	GGCC	AGACCCGCGA	GCACGCTCTG	TTGGCTTCA	CCCTCGGTGT	CAAGCAGATG	GTCGTCGCCA	TCAACAAGAT	GGACGACAAG	ACCGTCAAC-	[483]
045_EF1a_Trimastix_pyriformis	-----	GGCC	AGACCAACGA	GCACGCCCTG	CTCGCTTCA	CCCTCGGTGT	GAAGCAGGTT	ATCTCGTCTG	TCAACAAGAT	GGATGACGCC	AACGAGGACC	[418]
046_EF1a_Dinenympha_exilis	-----	GGTC	AAACTCGTGA	ACATGCAATT	CTGGCTTATA	CTCTGGGGTG	GAAGCAGATG	ATTGTTGGCG	TGAACAAGAT	GGACGACAAG	TCAGTGAAC-	[423]
047_EF1a_Monocercomonoides_sp.	-----	GGAC	AAACTCGTGA	GCACGCTCTT	CTTGCTTACA	CATTGGGTGT	GAAGCAGATG	ATCGTCTTGT	TCAACAAGAT	GGATGACAAA	TCTGTCAAC-	[483]
048_EF1a_Streblospio_trix	-----	GGAC	AAACACGTGA	ACACGCCCTT	CTTGCTTACA	CATTGGGTGT	GAAGCAGATG	ATTGTCCTTG	TAAACAAGAT	GGACGACAAA	TCAGTGAGC-	[432]
049_EF1a_Histomonas_meleagridis	-----	GGCC	AAACCCGTGA	GCACGCTCTT	CTTGCTTCA	CCCTTGGGTGT	CAAGCAAATG	ATTGTTGGTG	TTAACAAGAT	GGATGACAAG	ACAGTCAAC-	[483]
050_EF1a_Trichomonas_vaginalis	-----	GGCC	AGACACGTGA	ACACGCTCTT	CTTGCTTCA	CCCTTGGGTGT	CAAGCAGGTT	ATTGTCGCCG	TCAACAAGAT	GGATGACAAG	ACAGTCAAT-	[483]
051_EF1a_Hexamita_inflata	-----	GGTC	AAACTCGTGA	ACACGCCACT	CTTGCCAACA	CCCTTGGGTGT	CAAGCAATG	ATTGTTGCCG	TCAACAAGAT	GGACGACCCA	CAAGTCAAC-	[432]
052_EF1a_Spironucleus_vortens	-----	GGCC	AGACCCGCGA	GCACGCCACT	CTCGCCAACA	CCCTCGGTGT	CAAGCAATG	ATCTGTGCTA	TCAACAAGAT	GGACGACCCG	AACGTCAAC-	[432]
053_EF1a_Andalucia_incarcerata	-----	GGTC	AGACTCTGTA	GCACGCTCTT	CTTGCTTCA	CCCTTGGGTGT	TAAGCAGATG	ATCGTCGGTG	TGAACAAGAT	GGACGACAAG	TCCGTCAAC-	[432]
054_EF1a_Seculamonas_ecuadoriensis	-----	GGTC	AGACCCGCGA	GCACGCCCTG	CTCGCTTCA	CCCTCGGTGT	GAAGCAGATG	GCGCTGCCA	TCAACAAGAT	CGACGACAAG	TCGGTGAAC-	[483]
055_EF1a_Reclinomonas_americana	-----	GGCC	AGACCCGTGA	GCACGCTCTG	CTCGCTTCA	CCCTCGGTGT	GAAGCAGATG	CTGGTCGCCG	TCAACAAGAT	CGACGACAAG	TCTGTGAAC-	[483]
056_EF1a_Histiona_roides	-----	GGCC	AGACCCGTGA	GCACGCTTTG	CTCGCTTCA	CCCTCGGTGT	GAAGCAAATG	CGCGTCGGTG	TCAACAAGAT	GGACGACAAG	TCGGTCAAC-	[483]
057_EF1a_Jakoba_libera	-----	GGAC	AGACCCGTGA	GCACGCTCTC	CTCGCTTCA	CCCTGGGTGT	CAAGCAGATG	TGCGTGGGGA	TCAACAAGAT	CGACGACAAG	TCCGTCAAC-	[483]
058_EF1a_Stachyamoeba_lipophora	-----	GGTC	AGACCCGTGA	GCACGGTCTG	CTCGCTTCA	CCCTCGGTGT	CAAGCAGATG	GTCGTTGCTT	GCAACAAGAT	GGATGATACT	TCCGTCAAC-	[483]
059_EF1a_Naegleria_gruberi	-----	GGTC	AAACTAGAGA	ACACGCTTTG	CTCGCTTCA	CTTTGGGTAT	TAAGCAAATG	ATTGTCGCTA	TGAATAAGTT	TGATGATACT	TCTGTCAAT-	[417]
060_EF1a_Naegleria_andersoni	-----	GGTC	AAACTAGAGA	ACACGCTTTG	CTCGCTTCA	CTTTGGGTAT	TAAGCAAATG	ATTGTCGCTA	TGAACAAGTT	TCGATGATACT	TCTGTCTTCT-	[432]
061_EF1a_Acrasis_rosea	-----	GGAC	AAACCAAGGA	ACACGCCATG	TTGGCTTCA	CCCTTGGGTGT	CAAGCAAATG	ATCGTCTGCA	CCAACAAGAT	GGACGACAAG	TCAGTTCAA-	[432]
062_EF1a_Sawyeria_marylandensis	-----	GGAC	AAACTAGAGA	ACATGCTTTA	TTAGCATCAA	CTTGGGTGT	TAAATAAATG	TTAGTTATTG	TAAATAAATG	GGATGATACT	ACAAACAATG-	[483]
063_EF1a_Bodo_saltans	-----	GGCC	AGACCCGCGA	GCACGCCCTG	CTCGCTTCA	CCCTCGGCGT	CAAGCAGATG	GTCGTTGCTT	GCAACAAGAT	GGAC-----	TCCATCAAC-	[423]
064_EF1a_Rhynchobodo_sp.	-----	GGAC	AGACCCGCGA	GCACGCCCTC	CTCGCTTCA	CCCTCGGTGT	CAAGCAGATG	ATCGTTGCTT	GCAACAAGAT	GGACGACAAG	TCCGTCAAC-	[417]
065_EF1a_Rhynchomonas_nasuta	-----	GGCC	AGACCCGCGA	GCACGCCCTG	CTCGCTTCA	CGCTGCGGCT	CAAGCAGATG	ATCGTCGGTG	GCAACAAGAT	GGACGCCGAC	TCCGTCAAG-	[480]
066_EF1a_Dimastigella_trypaniformis	-----	GGCC	AGACCCGCGA	GCACGCCCTG	CTCGCTTCA	CTCTGGGCGT	CAAGCAGATG	ATCGTCTTCT	TCAACAAGAT	GGACGACAAG	TCGGTCAAC-	[438]

	450	460	470	480	490	500	510	520	530	540	550	
067_EF1a_Perkinsiella_amoebae	-----	GGCC	AGACCCGCGA	GCACGCGCTC	CTCGCCTTCA	CCCTCGGCGT	CAAGCAGATG	ATCGTCGCGG	TCAACAAGAT	GGACGACAAG	ACTGTCAAC	[408]
068_EF1a_Leishmania_major	-----	GGCC	AGACCCGCGA	GCACGCGCTG	CTTGCCCTTCA	CTCTTGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACGGTGACG	[483]
069_EF1a_Leishmania_infantum	-----	GGCC	AGACCCGCGA	GCACGCGCTG	CTTGCCCTTCA	C6CTTGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACCGTGACG	[483]
070_EF1a_Leishmania_braziliensis	-----	GGCC	AGACCCGCGA	GCACGCGCTG	CTCGCCTTCA	C6CTGGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACGGTGACG	[483]
071_EF1a_Leishmania_tarentolae	-----	GGGC	AGACCGCGGA	GCACGCGCTG	CTTGCCCTTCA	CTCTTGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACGGTGATG	[483]
072_EF1a "L. tarentolae" FJ807257	-----	GGGC	AGACACGCGA	GCACGCGCTT	CTCGCCTTCA	C6CTTGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACGGTGTCG	[408]
073_EF1a_Leptomonas_costaricensis	-----	GGCC	AGACCCGCGA	GCACGCGCTG	CTCGCCTTCA	C6CTGGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACCGTGACG	[231]
074_EF1a_Leptomonas_bifurcata	-----	GGCC	AGACCCGCGA	GCACGCGCTG	CTCGCCTTCA	C6CTGGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACCGTGACG	[243]
075_EF1a_Leptomonas_podlipaevi	-----	GGCC	AGACCCGCGA	GCACGCGCTG	CTGGCCTTCA	C6CTGGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACCGTGACG	[231]
076_EF1a_Herpetomonas_muscarum	-----	GGTC	AGACCCGCGA	GCACGCGCTC	CTCGCCTTCA	C6CTGGGCGT	GAAGCAGATG	ATCGTGTGCT	GCAACAAGAT	GGACGACAAG	ACCGTGCGC	[396]
077_EF1a_Herpetomonas_pessoai	-----	GGCC	AGACCCGCGA	GCACGCTCTC	CTCGCCTTCA	CCCTTGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACCGTGACG	[264]
078_EF1a_Blastocirithidia_culicis	-----	GGCC	AGACCCGCGA	GCACGCGCTG	CTCGCCTTCA	C6CTGGGCGT	GAAGCAGATG	ATCGTGTGCT	GCAACAAGAT	GGACGACAAG	ACCGTCAAC	[396]
079_EF1a_Trypanosoma_brucei	-----	GGAC	AGACCCGCGA	GCACGCGTTG	CTGGCCTTCA	CTTTGGGTGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACTGTGAAC	[483]
080_EF1a_T. brucei_equiperdum	-----	GGAC	AGACCCGCGA	GCACGCGTTG	CTGGCCTTCA	CTTTGGGTGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACTGTGAAC	[405]
081_EF1a_Trypanosoma_avium	-----	GGTC	AGACCCGCGA	GCACGCGCTC	CTCGCCTTCA	C6CTGGGTGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	ACCGTGAAC	[429]
082_EF1a_Trypanosoma_cruzi	-----	GGCC	AGACACGCGA	GCACGCGCTG	CTCGCCTTCA	C6CTGGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	TCGGTGAAC	[483]
083_EF1a_Trypanosoma_rangeli	-----	GGCC	AGACACGCGA	GCACGCGCTG	CTCGCCTTCA	C6CTGGGCGT	GAAGCAGATG	GTGGTGTGCT	GCAACAAGAT	GGACGACAAG	TCGGTGAAC	[483]
084_EF1a_Ploetitia_costatum	-----	GGTC	AGACCCGCTG	GCACGCTCTG	CTGGCCTTCA	CCCTTGGTGT	AAAGCAGATG	ATCGTTGCCA	CCAACAAGTT	CGATGACAAG	ACTGTCAAG	[480]
085_EF1a_Ploetitia_oblonga_cDNA	-----	GGTC	AGACCCGCTG	GCACGCTCTG	CTGGCCTTCA	CCCTTGGTGT	AAAGCAGATG	ATCGTTGCCA	CCAACAAGTT	CGATGACAAG	ACTGTCAAG	[480]
086_EF1a_Ploetitia_oblonga_gDNA	-----	GGGC	AGACCCGCTG	GCACGCGCTT	TTGGCCTTCA	CCTTGGGAGT	GAAGCAAAAT	ATCGTTGCCA	GCAACAAGAT	GGACGACAAG	ACCGTCAAG	[480]
087_EF1a_Entosiphon_sp.	-----	GGTC	AGACCCGCTG	GCACGCGCTT	CTCGCCTTCA	CCTTGGGTGT	GAAGCAGATG	ATCGTGGCTT	GCAACAAGAT	GGACGACAAG	ACAGTGAAG	[480]
088_EF1a_Entosiphon_sulcatum	-----	GGCC	AGACCCGCTG	GCACGCGCTT	CTTGCTTACA	CTTTGGGTGT	CAAGCAAAAT	ATCGTTGCCA	GCAACAAGAT	GGACGACAAG	ACCGTGAAG	[408]
089_EF1a_Peranema_tricophorum	-----	GGTC	AAACCCAGAGA	ACACGCTCTC	TTGGCCTACA	CACCTTGGGT	AAACCAAAAT	ATTGTTGCCG	TTAACAAGAT	GGACGACAAG	ACCGTGAAG	[480]
090_EF1a_P. tricophorum_FJ807243	-----	GGTC	AAACCCGCTG	ACACGCGCTT	TTGGCCTACA	CCTTGGGTGT	CAAGCAAAAT	ATTGTCGCTG	TCAACAAGAT	GGACGACAAG	ACCGTGAAG	[474]
091_EF1a_Colacium_elongatum	-----	GGTC	AGACCCGCTG	GCACGCGCTT	CTGGCCTTCA	CGTTGGGTGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	TGATGACAAG	ACTGTGAAG	[480]
092_EF1a_Colacium_mucronatum	-----	GGTC	AGACCCGCTG	GCACGCGCTT	CTGGCCTTCA	CCCTGGGGGT	GAAGCAGATG	ATCGTCGCCA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[426]
093_EF1a_Colacium_sp.	-----	GGCC	AGACCCGCTG	GCACGCTCTG	TTGGCCTACA	CCCTGGGTGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[480]
094_EF1a_Colacium_vesiculosum	-----	GGTC	AGACCCGCTG	GCACGCGCTT	CTGGCCTTCA	CCCTGGGGGT	GAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[438]
095_EF1a_Cryptoglena_pigra	-----	GGCC	AGACCCGCTG	GCACGCTCTG	CTGGCCTTCA	CCCTTGGTGT	GAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[480]
096_EF1a_Cryptoglena_skujai	-----	GGCC	AGACCCGCTG	GCATGCTCTG	CTGGCCTTCA	CCCTTGGTGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	TGATGACAAG	ACTGTGAAG	[480]
097_EF1a_Discoplistis_adunca	-----	GGGC	AGACCGCTG	GCACGCTCTG	CTGGCCTTCA	CACCTTGGGT	GAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	CGATGACAAG	TCTGTGATG	[480]
098_EF1a_Euglena_clara	-----	GGGC	AGACTCGTGA	GCATGCTCTG	CTGGCCTTCA	C6CTGGGTGT	GAAGCAAAAT	ATCGTCGCTA	CCAACAAGTT	CGACGACAAG	ACAGTGAAG	[360]
099_EF1a_Euglena_deses	-----	GGCC	AGACCCGCTG	ACACGCTCTG	CTTGCTTACA	CCCTTGGTGT	CAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	TGATGACAAG	ACAGTCAAA	[480]
100_EF1a_Euglena_exilis	-----	GGGC	AGACCCGCTG	GCACGCGCTT	CTTGCTTACA	CACCTGGGAGT	GAAGCAAAAT	ATCGTTGCCA	CCAACAAGTT	TGATGATAAA	ACGGTCAAG	[360]
101_EF1a_Euglena_geniculata	-----	GGTC	AGACCCGCTG	GCACGCTCTG	CTGGCCTTCA	CCCTTGGTGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[480]
102_EF1a_Euglena_gracilis	-----	GGAC	AGACTCGTGA	GCATGCTCTG	CTTGCTTACA	CCCTGGGTGT	GAAGCAAAAT	ATCGTTGCCA	CTAATAAGTT	CGACGACAAG	ACTGTGAAG	[480]
103_EF1a_E. gracilis_X16890	-----	GGAC	AGACACGCTG	ACACGCACTT	TTGGCATAAG	CCCTTGGTGT	GAAGCAAAAT	ATTGTTGCCA	CCAACAAGTT	CGATGACAAG	ACAGTGAAG	[483]
104_EF1a_E. gracilis_FJ807238	-----	GGAC	AGACTCGTGA	GCATGCTCTG	CTTGCTTACA	CCCTGGGTGT	GAAGCAAAAT	ATCGTTGCCA	CTAATAAGTT	CGACGACAAG	ACTGTGAAG	[483]
105_EF1a_Euglena_granulata	-----	GGCC	AGACACGCTG	GCACGCACTG	CTGGCCTTCA	CCCTTGGTGT	GAAGCAGATG	ATTGTTGCTA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[480]
106_EF1a_Euglena_hiemalis	-----	GGCC	AGACCCGCTG	GCACGCTTGT	CTTGCTTACA	CCCTTGGTGT	CAAGCAAAAT	ATTGTTGCCA	CCAACAAGTT	CGACGACAAG	ACCGTGAAG	[480]
107_EF1a_Euglena_longa	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
108_EF1a_Euglena_stellata	-----	GGTC	AAACCCGCTG	ACATGCTCTG	CTTGCTTACA	CTTTGGGAGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	TGATGACAAG	ACTGTGAAG	[360]
109_EF1a_Euglena_viridis	-----	GGTC	AGACTCGTGA	GCACGCGCTG	CTTGCTTACA	CCCTGGGAGT	TAAACAGATG	ATTGTTGCCA	CCAACAAGTT	TGATGACAAG	TCTGTGAAG	[438]
110_EF1a_Eutreptia_pertyi	-----	GGCC	AGACACGCTG	GCACGCGCTG	CTGGCTTACA	CCCTTGGTGT	GAAGCAGATG	ATTGTTGCTA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[480]
111_EF1a_Eutreptiella_pomquetensis	-----	-----	AGACCCGCTG	GCACGCTCTG	CTGGCCTTCA	CCCTTGGTGT	AAAGCAGATG	ATCGTTGCCA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[90]
112_EF1a_Lepocinlis_acus_var._major	-----	GGGC	AGACCGCTG	GCACGCTCTG	CTTGCTTACA	CCTTGGGTGT	AAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	TGATGACAAG	ACTGTGAAG	[480]
113_EF1a_Lepocinlis_ovum	-----	GGAC	AGACGCTGTA	GCACGCTCTG	TTGGCTTACA	CCTTGGGTGT	TAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	CGATGACAAG	ACAGTGAAG	[360]
114_EF1a_Lepocinlis_spyrogyroides_clone_38	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
115_EF1a_L. spyrogyroides_clones_1_&_4	-----	GGTC	AGACCCGCTG	GCACGCTCTT	CTGGCCTTCA	CTCTGGGTGT	GAAGCAGATG	ATCGTTGCCA	CCAACAAGTT	TGATGACAAG	ACCGTCAAA	[480]
116_EF1a_Lepocinlis_steinii	-----	GGGC	AGACGCTGTA	GCACGCTCTG	CTGGCTTACA	CCCTGGGGGT	CAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	TGACGACAAG	ACCGTCAAG	[360]
117_EF1a_Lepocinlis_triapteris	-----	GGGC	AGACGCTGTA	GCACGCTCTG	CTGGCTTACA	CCTTGGGAGT	GAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	TGATGACAAG	ACCGTGAAG	[480]
118_EF1a_Monomorphina_aenigmatica	-----	GGGC	AGACGCTGTA	GCACGCTCTG	CTGGCTTACA	CCCTTGGTGT	TAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	CGATGACAAG	ACCGTGAAG	[480]
119_EF1a_Monomorphina_pyrum	-----	GGTC	AGACCCGCTG	GCACGCTCTT	CTGGCCTTCA	CTCTGGGTGT	GAAGCAGATG	ATCGTTGCCA	CCAACAAGTT	TGATGACAAG	ACCGTCAAA	[480]
120_EF1a_Monomorphina_sp.	-----	GGGC	AAACCCGCTG	GCATGCTCTT	CTGGCCTTCA	CTCTGGGTGT	GAAGCAGATG	ATCGTCGCAA	CCAACAAGTT	TGATGACAAG	ACTGTGCC	[480]
121_EF1a_Phacus_orbicularis	-----	GGTC	AGACGCTGTA	GCACGCACTT	CTGGCTTACA	CCCTTGGTGT	CAAGCAGATG	ATTGTTGCAA	CCAATAAGTT	TGATGATAAG	ACAGTGAAG	[438]
122_EF1a_Phacus_similis	-----	GGGC	AGACGCTGTA	GCACGCTCTG	CTGGCTTACA	CCCTTGGTGT	TAAGCAGATG	ATTGTTGCAA	CCAACAAGTT	TGACGACAAG	ACCGTGAAG	[480]
123_EF1a_Strombomonas_acuminata	-----	GGGC	AGACCCGCTG	GCATGCTCTG	TTGGCCTTCA	CATTGGGTGT	GAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	TGATGACAAA	ACTGTGAAA	[483]
124_EF1a_Strombomonas_borsteniensis	-----	GGTC	AGACCCGCTG	GCACGCTCTG	CTGGCTTACA	CCCTGGGTGT	GAAGCAAAAT	ATTGTCGCTA	CCAACAAGTT	CGATGATAAG	ACTGTGAAG	[360]
125_EF1a_Strombomonas_napiformis	-----	GGTC	AGACCCGCTG	GCACGCTCTG	CTTGCTTACA	CCTTGGGTGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	TGATGATAAG	ACAGTGAAG	[480]
126_EF1a_Strombomonas_ovalis	-----	GGTC	AGACACGCTG	GCACGCGCTG	CTGGCCTTCA	CCCTGGGTGT	GAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	CGATGACAAG	ACTGTGAAG	[480]
127_EF1a_Strombomonas_triquetra	-----	GGGC	AGACTCGTGA	GCATGCTCTG	CTGGCTTACA	C6CTGGGTGT	GAAGCAAAAT	ATCGTTGCCA	CCAACAAGTT	TGATGACAGG	ACTGTGAAG	[480]
128_EF1a_Trachelomonas_conradi	-----	GGAC	AGACCCGCTG	GCACGCTCTG	CTGGCATAAG	C6CTGGGTGT	GAAGCAGATG	ATTGTCGCCA	CCAACAAGTT	TGATGACAAG	ACCGTGAAG	[480]
129_EF1a_Trachelomonas_echinata	-----	GGGC	AAACTCGTGA	ACATGCTCTG	CTTGCTTACA	CATTGGGTGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	TGATGACAAG	ACCGTGAAG	[480]
130_EF1a_Trachelomonas_lefevrei	-----	GGCC	AGACTCGTGA	GCACGCTCTG	CTTGCTTACA	CACCTGGGTGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	CGACGACAAG	ACCGTGAAG	[480]
131_EF1a_Trachelomonas_plantonica	-----	GGTC	AGACTCGTGA	GCACGCTCTG	CTTGCTTACA	CACCTGGGTGT	GAAGCAGATG	ATTGTTGCCA	CCAACAAGTT	CGACGACAAG	ACCGTGAAG	[360]
132_EF1a_Trachelomonas_similis	-----	GGTC	AGACACGCTG	GCATGCTTGT	CTGCTTACA	CACCTTGGTGT	GAAGCAAAAT	ATTGTTGCAA	CCAACAAGTT	CGACGACAAG	ACCGTGAAG	[360]

		450	460	470	480	490	500	510	520	530	540	550		
133	EFL1a	Trachelomonas_volvocina	-----GGGC	AGACCCGAGA	ACATGCTCTG	CTTGCTTACA	CTCTTGGTGT	GAAGCAGATG	ATTGTGGCCA	CCAACAAGTT	TGATGACAAG	ACAGTCAAG-	[480]	
134	EFL	environmental_sample_EU810232	GGCCGGAGAG	GTGCAAGGCC	AGACCAAGCA	GCACGCTCTG	TGATCAAACT	TGCTGGGAGT	CAAGCAGCTC	TATGTCGGAG	CCAACAAGAT	GGACTGCGGT	ACCGCCAAAG	[504]
135	EFL	Creolimax_fragrantissima	GGCCGGTGAG	GTCCAGGGAC	AGACTCGTCA	GCACGCTTTG	TTGATCAACC	TTTTGGGTGT	TAAGCAGCTC	ATTGTTGGTG	TAAACAAGAT	GGATTGCGAT	GTGCGCAAG	[432]
136	EFL	Sphaeroforma_arctica	GGCCGGAGAG	GTCCAAGGGAC	AGACTCGTCA	GCACGCTCTT	CTTATCAACC	TTCTAGGTGT	TAAGCAGATTG	AYYNNNGNTN	TNAACAAGAT	GGATTGCGAT	GTGCGCAAG	[504]
137	EFL	Monosiga_brevicollis	GGCCGGTGAG	GTCCAAGGGCC	AGACCCGGCA	GCACGCGCTG	CTCATCAACC	TGCTCGGCGT	CAAGCAAGCT	ATTGTTGGCT	GCAACAAGAT	GGATTGCGAG	GTGCGCGGC	[504]
138	EFL	Xenopus_laevis	GGCCGGTGAG	GTCCAAGGGCC	AGACCCGGCA	GCACGCGCTG	CTCATCAACC	TGCTCGGCGT	CAAGCAGCTC	ATCGTCGGCG	TCAACAAGAT	GGACTGCGAT	GTGCGCAAG	[504]
139	EFL	Spizellomyces_punctatus	GGCCGGTGAA	ATTCAGGGCC	AGACTCGTCA	GCACGCTCTG	CTTCTCAACC	TTCTCGGTGT	GAAGCAGCTT	GTTGTTGGTG	TGAACAAGAT	GGACTCTGAC	CCCGCTGGTC	[505]
140	EFL	Gaertneriomyces_semiglobiferous	GGCTGGTGAA	ATCCAAGGGTC	AGACCCGGCA	GCACGCTCTG	CTCCTCAACT	TGCTCGGTGT	CAAGCAGCTC	GTCGTTGGTG	TCAACAAGAT	GGACTCTCGAC	CCTGCTGGTC	[208]
141	EFL	Allomyces_macrogynus	GGCCGGTGAA	ATCCAAGGGCC	AGACCGGCCA	GCACGCTCTG	CTGCTCAACC	TGCTCGGCTG	CAAGCAGCTC	GTTGTCGGTG	TCAACAAGAT	GGACTCTGGAC	CCGCGCGGTC	[505]
142	EFL	Blastocladiella_emersonii	GGCCGGTGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCTCTG	CTGCTCAACC	TGCTCGGTGT	CAAGCAGCTC	GTCGTCGGTG	TCAACAAGAT	GGACTCTCGAC	CCCGCTGGTC	[505]
143	EFL	Catenophylctis_sp.	GGCCGGTGAA	ATTTCAAGGGC	AAACCCGGCA	GCACGCGCGT	CTTCTCAACT	TGCTCGGTGT	CAAGCAGCTC	GTTGTTGGTG	TCAACAAGAT	GGACTCTGAC	CCCGCTGGTC	[208]
144	EFL	Coelomomyces_stegomyiae	GTCCGGAGAA	ATTTCAAGGAC	AAACCCGTCA	ACATGCCCCT	TTATTAACCT	TGTTGGGTGT	TAAGCAATTG	ATTGTCGGTA	TCAACAAGAT	GGACACTGAC	ACTGCTGGTC	[208]
145	EFL	Entomophthora_muscae	GGCCGGAGAT	ATCCAAGGGAC	AGACCCGGCA	GCATGCCCCT	CTTCTCAACT	TTCTCGGTGT	TAAGCAGCTC	ATCATTTGGTA	TAAACAAGAT	GGACTGTGAC	ACCGCTGGAA	[207]
146	EFL	Basidiobolus_ranarum	GGCCGGAGAT	ATCCAAGGGTC	AAACTCGCCA	ACATGCTCTG	CTTCTTAACC	TTCTCGGTGT	TAAGCAAGAT	ATCATYGGTG	TAAACAAGAT	GGATTCCGAT	ACCGCTGGT	[207]
147	EFL	Conidiobolus_coronatus	GGCCGGTGAA	ATCCAAGGGTC	AAACCCGGCA	GCACGCTCTG	CTTCTTAACC	TACTCGGTGT	TAAGCAACTT	ATCATTTGGTG	TCAACAAGAT	GGATTCCGAT	GTGCTCAA	[207]
148	EFL	Tetraselmis_tetrashele	GGCCGGTGAG	GTGCAAGGGC	AGACCCGGCA	GCACGCGCGT	CTGATCAACC	TGCTGGGCGT	GAAGCAGCTC	ATCATCGGTA	TCAACAAGAT	GGACTCTGAT	GTGCTGGC	[453]
149	EFL	Mesostigma_viride	GGCCGGTGAG	GTCCAAGGGCC	AGACCCGGCA	GCACGCTCTG	CTGATCAACC	TGCTGGGTGT	GAAGCAGCTT	TGCTCGGTG	TCAACAAGAT	GGACTCCGAT	ACCGCTGGC	[507]
150	EFL	Ostreococcus_tauri	GGCTGGTGAA	GTGCAAGGGC	AAACCCGGCA	GCACGCGCGT	CTTCTCAACT	TGCTCGGCGT	CAAGCAGCTC	ATCATCGGTA	TCAACAAGAT	GGACTGCGAT	ATGGCTGGC	[504]
151	EFL	Nephrolepisma_olivacea	GGCCGGTGAG	GTGCAAGGGC	AGACCCGGCA	GCACGCTCTG	CTGCTGAACC	TGTTGGGCGT	GAAGCAGCTG	ATTGTCGGTG	TGAACAAGAT	GGACTCCGAC	ATTGCGGCG	[480]
152	EFL	Micromonas_pusilla	GGCCGGTGAG	GTCCAAGGGCC	AGACCCGGCA	GCACGCGCGC	CTCATCAACC	TTCTCGGCGT	CAAGCAGCTC	ATCGTCGGCG	TCAACAAGAT	GGACTGCGAC	ACCGCGGGC	[453]
153	EFL	Parachlorella_kessleri	GGCCGGCGAG	ATCCAAGGGCC	AGACCCGGCA	GCATGCTCTG	CTCATCAACC	TGCTGGGCGT	GAAGCAGCTC	ATCGTCGGCG	TGAACAAGAT	GGATTCTGAC	ACTGCTGGC	[462]
154	EFL	Helicosporidium_sp.	GGCCGGCGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCGCGC	CTGCTCAACC	TGCTGGGCGT	GAAGCAGCTC	ATCGTCGGTG	TGAACAAGAT	GGACTCTGAC	GTGCGCAAG	[504]
155	EFL	Gonium_pectorale	GGCCGGTGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCGCGC	CTGATCAACC	TGCTCGGCGT	GAAGCAGCTG	ATTGTCGGTG	TCAACAAGAT	GGACTCCGAC	ACTGCGGCG	[504]
156	EFL	Chlamydomonas_reinhardtii	GGCCGGTGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCTCTG	CTGATCAACC	TGCTGGGTGT	CAAGCAGCTG	ATTGTTGGTG	TGAACAAGAT	GGACTCCGAC	ACTGCGGCG	[504]
157	EFL	Scenedesmus_obliquus	GGCCGGTGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCGCGC	CTGATCAACC	TGCTGGGCGT	GAAGCAGCTG	ATTGTCGGTG	TGAACAAGAT	GGACACGCGAT	GTGGCGGCG	[504]
158	EFL	Oryza_sativa	GGCCGGTGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCGCGC	CTGATCAACC	TGCTGGGTGT	GAAGCAGCTG	ATTGTCGGTG	TGAACAAGAT	GGACTCTGAT	GTGGCTGGC	[504]
159	EFL	Chlorococcum_sp.	AGCCAAATAG	GTCCAAGGGCC	AGACCAAGCA	GCACGCGCGC	CTTATTAACC	TTCTTGGTGT	CGCCGAGCTC	TGCAATTGGTG	TCAACAAGAT	GGATTCTGAC	ACTGCTGGT	[453]
160	EFL	Ochlochaete_hystrix	AAGCGGTGAG	ATTTCAAGGGC	AGACCCGGCA	GCACGCGCGC	CTGATCAACC	TGCTGGGTGT	GAAGCAGCTG	GTTGTGGGTA	TCAACAAGAT	GGACTCCGAC	ACTGCTGGC	[462]
161	EFL	Ulva_intestinalis	GGCCAAATAG	GTCCAAGGGCC	AGACCCGGCA	GCATGCTCTG	CTGCTCAACC	TTCTCGGTGT	CGCTGAGCTG	TGTTTGGTG	TCAACAAGAT	GGATGCGGAC	ACTGCTGGC	[453]
162	EFL	Urospora_sp.	GGCCAAAGAG	GTCCAAGGGCC	AGACCGGCCA	GCACGCGCGC	TGCTGCTGCG	TCTCGGCGT	CGCCGAGCTC	TGCGTGGGTG	TCAACAAGAT	GGACTGCGAC	GTGCGGCGC	[306]
163	EFL	Gymnochlorella_stellata_ppc	GGTGGGCGAG	GTCCAAGGGCC	AAACCAAGCA	GCACGCAAGA	CTTCTCAACT	TGCTCGGAGT	GAACAAATTG	ATATGTGCTG	TGAACAAGAT	GGATTGTGAC	CAGCTAAG	[540]
164	EFL	Lotharella_amoebiformis_ppc	ACTAGGAGAG	GTCCAAGGGTC	AAACCGGCCA	GCACGCAAGT	CTTCTCAACC	TCTCGGGGTG	CAAGCAGCTG	ATTGTTGGCA	TAAACAAGAT	GGACTGCGAC	CAGGCTAAA	[285]
165	EFL	Chlorarachnion_reptans_ppc	GGTGGGTGAG	GTGCAAGGGC	AAACAAGGCA	TCACGCAACG	CTTCTGAACC	TTCTTGGGCT	TAAGCAGCTC	ATTGTTGGCTG	TCAACAAGAT	GGACTGTGAG	CAGGCTAAA	[312]
166	EFL	Chondrus_crispus	GGCCGGTGAG	GTCCAAGGGCC	AGACCCGGCA	GCATGCTCTG	CTTCTCAACC	TTTTGGGCGK	CAMGMRGCTA	ATTGTCGGTG	TGAACAARAT	GGATTGCGAC	CTGCGCAAG	[504]
167	EFL	Gracilaria_changii	GGCTGGTCA	GTGCAAGGGC	AAACTCGCCA	GCATGCGCGT	TGCTCAACC	TGCTGGGTGT	CAACAAGATC	ATCGTAGGTT	TCAACAAGAT	GGATTGTGAT	CTGCGCAAG	[504]
168	EFL	Cryptomonas_ovata	GGCCGGTGAG	ATCCAAGGGCC	AGACTCGTCA	GCACGCTCTG	ATCCTCAACC	TTCTCGGCGT	GAAGCAGCTC	GTTGTCGGTG	TGAACAAGAT	GGACTCCGAC	ACTGCGGCG	[459]
169	EFL	Guillardia_theta	GGCTGGAGAG	ATCCAAGGGAC	AGACCCGGCA	GCACGCTCTG	ATCCTCAACC	TTCTCGGTAT	CAAGCAGCTC	ATCGTCGGCA	TCAACAAGAT	GGACTCCGAC	ACTGCTGGA	[501]
170	EFL	Rhodomonas_salina	GGCCGGAGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCGCGT	ATCCTCAACC	TTCTCGGTGT	GAAGCAGCTC	ATCGTCGGCA	TCAACAAGAT	GGATTGTGAC	GTGCGGCGC	[501]
171	EFL	Pavlova_lutheri	GGCCGGCGAG	GTGCAAGGGC	AGACCGGCCA	GCACGCGCGC	CTCCTCAACC	TGCTCGGCGT	GAAGCAGCTC	ATCGTCGGCG	TCAACAAGAT	GGACTCGGAC	CCGCGGGGCG	[502]
172	EFL	Chrysochromulina_sp.	GGCCGGTGAG	GTCCAAGGGCC	AGTCCGGCCA	GCACGCGCGC	CTCATCAACC	TGCTCGGTGT	CAAGCAGCTG	CTGATTGGCG	TAAACAAGAT	GGACACCGAC	ACCGCTGGCC	[322]
173	EFL	Isochrysis_galbana	GGCCGGCGAG	GTCCAAGGGCC	AGTCCGGCCA	GCATGCGCGC	CTCATCAACC	TGCTCGGCGT	CAAGCAGCTC	ATCATTTGGCG	TCAACAAGAT	GGACACCGAT	GTGGCGGCG	[504]
174	EFL	Emiliana_huxleyi	GGCCGGCGAG	GTGCAAGGGC	AGAGCCGGCA	GCACGCGCGC	CTCCTCAACC	TGTTGGGCGT	CAAGCAGCTG	CTGCTGGCG	TCAACAAGAT	GGACACCGAC	ACCGCGGGCG	[502]
175	EFL	Achnanthes_kuwaitensis	GATGGGAGTT	GAGGAGGGAC	AGACCCGGCA	GCACGCTCTG	CTTCTGACT	TGCTCGGTGT	GGAGCAGATC	ATTGTCGGAA	TCAACAAGAT	GGAC	GGCTGCAAC	[435]
176	EFL	Asterionella_glacialis	GTTGGGAGTT	GAGGAAGGTC	AAACCCGTCA	GCACGCGCGT	CTTTGTACT	TGCTCGGTGT	TGAGCAAAATC	ATTGTTGGAA	TCAACAAGAT	GGAC	GGCTGCGAA	[435]
177	EFL	Detonula_confervacea	GGCCGGAGTT	GAGGAGGGTC	AGACCCGTCA	GCACGCGCGT	CTTCTTTACC	TTTTGGGAGT	TGACCAAGATC	ATTGTTGGCA	TCAACAAGAT	GGAC	GGCTGTCAAG	[435]
178	EFL	Thalassiosira_pseudonana	GCTTGGAGTT	GAGGAGGGAC	AGACCCGTCA	GCACGCGCGT	CTTCTTTACC	TTCTTTGAGT	TGAGCAAAATC	ATTGTTGGAA	TCAACAAGAT	GGAC	GGCTGCAAC	[498]
179	EFL	Perkinsus_marinus	GGCTGGTGAG	ATCCAAGGGTC	AAACTCGTCA	ACATGCCCCT	CTTATTAATT	TGTTGGGCGT	CAAGCAACTA	ATCGTCGGCG	TGAACAAGAT	GGACTGTGAT	GTTGCGGCG	[513]
180	EFL	Oxyrrhis_marina	GGCTGGAGAG	ATCCAAGGGAC	AGACCCGGCA	GCACCTCCCG	TTGATCAACT	TGCTGGGAGT	GAAGCAGATC	ATCATCGGAG	TGAACAAGAT	GGATTGCGAT	TTGCTGGG	[510]
181	EFL	Karlodinium_micrum	GGCTGGTGAG	ATCCAAGGGCC	AAACTCGACA	GCACCTCAGG	TTAATCAACT	TGCTTGGTGT	CAAGCAGATC	GCAATCGGCG	TAAACAAGAT	GGACTGCGAT	ACTGCTGGC	[510]
182	EFL	Amphidinium_carterae	GGCTGGTGAG	ATCCAAGGGCC	AGACCAAGCA	GCACCTCTCT	CTCATCAACT	TGTTGGGCGT	GAAGCAAAATC	TGCATTGTTG	TGAACAACAT	GGATTGCGAC	ACCGCTGGA	[510]
183	EFL	Heterocapsa_triquetra	GGCGGGTGAG	ATCCAAGGGCC	AGACTCGTCA	GCACCTCCCT	CTCATCAACC	TGCTGGGCGT	GAAGCAGATC	GCCATTGGTG	TGAACAAGAT	GGACTGCGAC	ACCGCGGGC	[510]
184	EFL	Planoglabratella_operculis	GAAAGGAGAG	ATCCAAGGGAC	AGACCCGGCA	GCATGCTCTG	TTGTCGCACT	TGTTGGGCACT	TGAGCAATTG	ATTGTTGGGCA	TCAACAAGAT	GGATGAACCA	TCGGTCAAT	[342]
185	EFL	Reticulomyxa_filosa	AAAAGGAGAA	ATTTCAAGGAC	AAACCCGGCA	ACATGCTCTG	TTGTCGCACT	TGTTGGGAACT	TGAACAATTG	ATAGTCGGTG	TCAACAAGAT	GGATGATCCT	TACGTCAAC	[432]
186	EFL	Lotharella_amoebiformis	GGCCAACTCC	GTCAGGGGTC	AAACCCGTCA	GCACGCGGAG	CTTACGAACC	TTCTTGGTAT	CAAGCAGATG	ATTGTTGGCG	TCAACAAGAT	GGACGACAAAG	ACCGTCAAC	[549]
187	EFL	Gymnochlorella_stellata	TGCTAACTCC	GAAAAAGGGAC	AAACTCGTCA	GCACGCTGAG	CTTACCAACC	TGCTCGGAAAT	TAAGCAAAATC	ATCGTCGGAG	TCAACAAGAT	GGACGAAAAAG	TCTGTTCAAG	[465]
188	EFL	Chlorarachnion_reptans	GGCTGGATGCT	CAGAAAGGGTC	AAACCCGGCA	GCATGCTGAG	CTCACCAAGC	TTCTCGGTAT	TCAACAAGAT	ATTGTTGGCG	TTAATAAGAT	GGACGACGAG	TCTGTTCAAG	[432]
189	EFL	Diplonema_ambulator	-----GGGGAG	ATCCAAGGGCC	AGACCAAGCA	GCACGCCCCT	CTGATCAACC	TGCTCGGTGT	GAAGCAGCTC	ATCGTCGGGA	TCAACAAGAT	GGACTGCGAC	GTGGCTCAG	[105]
190	EFL	Diplonema_papillatum	GGCCGGAGAG	ATCCAAGGGCC	AGACCCGTCA	GCACCTCGCT	CTGATCAACT	TGTTGGGAGT	GAAGCAGTTG	ATCGTCGGTG	TGAACAAGAT	GGACTGCGAC	GTGCTCAG	[477]
191	EFL	Rhynchonopas_euleeides	GGCCGGTGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCGCGT	CTGATCAACC	TGCTGGGTGT	GAAGCAGCTG	GTCGTTGGTG	TGAACAAGAT	GGACTGCGAC	GTGCGCCAG	[117]
192	EFL	Petalomonas_cantuscyni	GTCCGGCGAG	ATCCAAGGGCC	AGACCCGGCA	GCACGCGCGT	CTCCTCAACC	TGCTCGGCGT	GAAGCAGCTC	ATCGTTGGCG	TCAACAAGAT	GGACACCGAC	GTGCGCAAG	[348]
193	EFL	Trypanoplasma_borrelli	GACYGGTGAG	GTCCCGGGCC	AGACCCGGYAA	GCACGCGCGC	CTTCTGAAYC	TGCTYGGYGT	GAAGCAGCTG	ATTGTCGGYG	TGAACAAGAT	GGACAGYCTC	GACTCYGACG	[514]
194	EFL	Neobodo_salien	GACCCGGTGAG	GTTCCGGGCG	AGACCCGGCAA	GCACGCGCGC	CTGCTCAACC	TGCTCGGCACT	CAAGCAGCTG	ATCGTCGGCA	TCAACAAGAT	GGACTGCTCC	GTGCAACAG	[463]

	560	570	580	590	600	610	620	630	640	650	660	
001_aEF1a_Desulfurococcus_mobilis	-----	-----	---TACAGTG	AGAAAGAGATA	CAATGAGATC	AAGGAGATCC	TGGGGAAGTT	CCTGAAGGGA	CTC---GGCT	AC-----	-----	[552]
002_aEF1a_Sulfolobus_solfataricus	-----	-----	---TATGATG	AAAAACGTTA	TAAGGAGATT	GTTGACCAAG	TCAGTAAGTT	TATGAAGAAT	TAT---GGAT	TT-----	-----	[546]
003_aEF1a_Thermoplasma_acidophilum	-----	-----	---tAcAgTg	AgAaGaGgTa	CaAcGaGgTc	AaGgCcGaTg	CtGaGaAaCt	CcTcAgGtCc	AtA---GgTt	Tc-----	-----	[534]
004_aEF1a_Halobacterium_sp.	-----	-----	---TACGACG	AGTCCAATA	CAACGAAGTG	GTCTCCGGCG	TCAAGGACCT	CTTCGGGCG	GTC---GGCT	TC-----	-----	[519]
005_aEF1a_Methanocaldococcus_jannaschii	-----	-----	---TACAGCC	AAGAAGAATA	CGAAAAAATG	AAAAAGATGT	TATCAGAGCA	GTTATTAATA	GTCTTAGGTT	AC-----	-----	[534]
006_EF1a_Acanthamoeba_culbertsoni	-----	-----	---TGGGCCG	AGAACCCTTA	CAACGAGATC	CAGCGTGAGG	TCTCGGGCTA	CCTGAAGAAG	GTC---GGCT	AC-----	-----	[516]
007_EF1a_Neoparamoeba_branchipila	-----	-----	---TACTCCG	AGGACCGTTA	CACCGAGATC	AAGAAGGAGG	TCTCCGGTTT	CTTGAAGAAG	ATC---GGGT	AC-----	-----	[501]
008_EF1a_Planoprotostelium_aurantium	-----	-----	---TACTCCG	AGGCTCGTTT	CAACGAAATC	AGAAGTGAAG	TTTCTCTTTT	CATCAAGAAG	ATC---GGAT	AC-----	-----	[498]
009_EF1a_Physarum_polycephalum	-----	-----	---TGAGGCC	AGGCCCGTTA	CGATGAAATC	GTGAAGGAAA	CCTCTCTCTT	CGTCAAGAAG	ATC---GGAT	AC-----	-----	[498]
010_EF1a_Amoebidium_parasiticum	-----	-----	---TTCGCC	AGGATCGCTT	CAACGAAATC	GTGAACGAGG	TGTCCAATA	CATCAAGAAG	ATC---GGCT	AC-----	-----	[513]
011_EF1a_Capsaspora_owczaraki	-----	-----	---TTCGCTG	AGGAGCGTTA	CAACGAAATC	GTCACTGAGG	TCTCGAATA	CATCAAGAAG	ATC---GGCT	AC-----	-----	[543]
012_EF1a_Monosiga_ovata	-----	-----	---TACAGCG	AGAGCCGCTT	CAACGAGATC	AAGACTGAGG	TCTCGACCTA	CATCAAGAAG	ATC---GGCT	AC-----	-----	[522]
013_EF1a_Salpingoeca_amphoridium	-----	-----	---TACAGCG	AGGCCCGCTT	CAACGAGATC	AAGACTGAGG	TGTCCAATT	CATCAAGAAG	GTC---GGCT	TC-----	-----	[516]
014_EF1a_Ceanorhabditis_elegans	-----	-----	---TTCTCTG	AGGCCCGTTT	CACCGAGATC	ACCAACGAGG	TTTCTGGATT	CATCAAGAAG	ATC---GGAT	AC-----	-----	[549]
015_EF1a_Drosophila_melanogaster	-----	-----	---TACAGCG	AGGCCCGTTA	TGAGGAAATC	AAGAAGGAAG	TGCTCTCTTA	CATCAAGAAG	ATC---GGCT	AC-----	-----	[549]
016_EF1a_Ornithorhynchus_anatinus	-----	-----	---TACAGCG	AAGAAGGTTA	CGATGAAATT	GTCAGGAGG	TGACGGCTTA	CATCAAGAAG	ATT---GGCT	AC-----	-----	[549]
017_EF1a_Xenopus_laevis	-----	-----	---TTTAGCC	AGAAAAGGTT	TGAAGAAATT	ACTAAGGAAG	TCAGTGCTTA	CATTAAAGAAG	ATT---GGCT	AC-----	-----	[549]
018_EF1a_Saccharomyces_cerevisiae	-----	-----	---TGGGACG	AATCCAGATT	CCAAGAAATT	GTCAAGGAAA	CCTCCAATT	TATCAAGAAG	GTT---GGTT	AC-----	-----	[543]
019_EF1a_Neurospora_crassa	-----	-----	---TGGTCTC	AGACTCGTTT	CGAGGAGATC	ATCAAGGAGA	CCAAGAACTT	CATCAAGAAG	GTT---GGCT	AC-----	-----	[546]
020_EF1a_Cryptococcus_neoformans	-----	-----	---TGGTCCG	AGGACCGATT	CAACGAAATC	GTCAAGGAGA	CCAACGGTTT	CATCAAGAAG	GTT---GGTT	AC-----	-----	[543]
021_EF1a_Schizophyllum_comune	-----	-----	---TGGAGCG	AGGACCGTTT	CAACGAAATC	GTCAAGGAGA	CCTCCACCTT	CATCAAGAAG	GTC---GGCT	AC-----	-----	[543]
022_EF1a_Monoblepharis_insignis	-----	-----	---TACTCGG	AGGAGCGCTT	CAACGAAATT	GTTCTGTAGG	TGTGCAACTT	CGTTAAGAAG	GTT---GGTT	AC-----	-----	[429]
023_EF1a_Rhizophlyctis_harderi	-----	-----	---TGGTCTG	AGGCTCGTTT	CAATGAAATT	GTCAAGGAAT	TGTCTCTTTT	CATCAAGAAG	GTT---GGTT	TC-----	-----	[429]
024_EF1a_Harpochytrium_sp.	-----	-----	---TGGTCCG	AGGAGCGTTA	CAACGAGATC	GTGAAGGAGG	TGTCTCGTA	CATCAAGAAG	GTC---GGAT	AC-----	-----	[381]
025_EF1a_Absidia_glauca	-----	-----	---TGGTCTG	AACAACGCTT	CAACGAAATC	ATCAAGGAAAG	TGTCTGGCTT	CATCAAGAAG	ATC---GGTT	TC-----	-----	[543]
026_EF1a_Smittium_simulii	-----	-----	---TACTCTG	AAGAACGTTT	CACTGAAATT	ATTAAGGAAG	TATCTAACTT	TATCAAAAAG	GTT---GGTT	AC-----	-----	[516]
027_EF1a_Chara_australis	-----	-----	---TACTCGG	AGAACAGGTA	CAACGAAATT	AAGAAGGAGG	TGTCCACCTA	TCTGAAGAGA	GTC---GGGT	AC-----	-----	[498]
028_EF1a_Physcomitrella_patens	-----	-----	---TGACGCA	AGGCCCGATA	TGAGGAGATC	TGAAGGAGG	TGTCCGCGTA	CCTGAAGAAG	GTG---GGGT	AC-----	-----	[549]
029_EF1a_Acetabularia_acetabulum	-----	-----	---TACAGCG	ATGCCAGATA	CAATGAGATC	AAGACGGAAG	TTTCCAAGTA	CCTCAAAAAG	GTG---GGGT	AC-----	-----	[498]
030_EF1a_Chaetomorpha_coliformis	-----	-----	---TACAGTG	ATAAGCGTTA	CGAGGAAATC	ATGAAGGAAG	TGAAGAACTA	CATTAAAGAAG	GTT---GGTT	AC-----	-----	[348]
031_EF1a_Arabidopsis_thaliana	-----	-----	---TACTCCA	AGGCTAGGTA	CGATGAAATC	ATCAAGGAGG	TGTCTTCTTA	CCTGAAGAAG	GTC---GGAT	AC-----	-----	[549]
032_EF1a_Oryza_sativa	-----	-----	---TACTCCA	AGGCCCGTTA	TGATGAAATC	GTGAAGGAAG	TCTCATCTTA	CCTGAAGAAG	GTC---GGCT	AC-----	-----	[549]
033_EF1a_Cyanidioschyzon_merolae	-----	-----	---TGGTCTC	AGTCGCGTTA	TGATGAAATC	ACCAAAGAGT	TGAACCTGTA	TTTGAAGAAG	GTT---GGCT	AC-----	-----	[549]
034_EF1a_Porphyrha_yezoensis	-----	-----	---TGGTCTG	AGGACCGCTA	CGAGGAGGTT	TCCAAGGAGA	TGGACCTGTA	CCTGAAGAAG	GTC---GGGT	AC-----	-----	[549]
035_EF1a_Plasmodium_falciparum	-----	-----	---TATTCTG	AAGACAGATA	TGAAGAAATT	AAAAAGGAAG	TTAAGGATT	CTTGAAGAAG	GTA---GGAT	AC-----	-----	[543]
036_EF1a_Cryptosporidium_parvum	-----	-----	---TACAACG	AATCTCGTTT	TGATGAAATC	TTCAACGAAG	TTGATGGTTA	CCTCAAGAAG	GTT---GGTT	AC-----	-----	[537]
037_EF1a_Tetrahymena_pyriformis	-----	-----	---TTCTCTG	AAGAAGGATA	TCAAGAAATC	AAGAAGGAAT	TGTCCGACTA	CTTGAAGAAG	GTT---GGTT	AC-----	-----	[552]
038_EF1a_Stentor_coeruleus	-----	-----	---TGGTCAC	AGGCCAGATA	CGATGAAATC	AAGAATGAAG	TTCCCAATA	CTTGAAGAAA	GTT---GGCT	AC-----	-----	[504]
039_EF1a_Phytophthora_parasitica	-----	-----	---TACGGCC	AGGCCCGTTA	CGAGGAGATC	AAGAACGAGG	TATCCACGTA	CCTGAAGAAG	GTC---GGCT	AC-----	-----	[549]
040_EF1a_Mallomonas_rasilis	-----	-----	---TACTCTG	AGCAGAGATA	CAATGATATT	AAGGAGGAAG	TTTCCACCTA	CCTCAAGAAG	GTC---GGTT	AC-----	-----	[510]
041_EF1a_Phaeodactylum_tricornutum	-----	-----	---TACGGCC	AAGATCGTTA	CACCGAAATT	AAGAATGAAG	TCTCTGCCTA	CCTCAAGAAG	GTT---GGCT	AC-----	-----	[549]
042_EF1a_Thalassiosira_pseudonana	-----	-----	---TATTCTG	AGGATCGCTT	CAACGAGATC	AAAGACGAGG	TATCGCGCTA	CTTGACTCAT	CTC---GGCT	AC-----	-----	[549]
043_EF1a_Malawimonas_jakobiformis	-----	-----	---TTCTCCC	AGGATCGCTT	CGAGGAGATC	GTCAAGGAGG	TGTCCAGCTT	CTTGAAGAAG	ACC---GGCT	AC-----	-----	[480]
044_EF1a_Malawimonas_californiana	-----	-----	---TGGGGCC	AGGACCGCTA	TGAGGAGATC	AAGAAGGAGG	TGTCCAATA	CGTCAAGAAG	ATT---GGCT	AC-----	-----	[549]
045_EF1a_Trimastix_pyriformis	CCAAGACCGG	TGAGAACAAAG	CCCTTCTCTT	GGCGCCGTTT	CGATGAGATC	AAGGACAACG	TCTTCATCCG	CTTGAAGAAG	ATC---GGCT	AC-----	-----	[507]
046_EF1a_Dinenympha_exilis	-----	-----	---TGGGCAG	AGTCACTGTA	TAATGAAATC	AAGACCGAAA	TGAGGACATA	CTTGAAGAAG	ATT---GGGT	AC-----	-----	[489]
047_EF1a_Monocercomonoides_sp.	-----	-----	---TACAGCG	AAGCTCGTTT	CAATGAAATC	AAGGAGGAAA	TGGCAACTA	TCTGAAGAAG	ATT---GGAT	AC-----	-----	[549]
048_EF1a_Streblospio_mexicana	-----	-----	---TTTTCTG	AAGCAAGATA	CAACGAAATT	GTAGGAGAGA	TGAGGAATA	CTTGAAGAAG	ATT---GGAT	AT-----	-----	[498]
049_EF1a_Histomonas_meleagridis	-----	-----	---TACTCCA	AGCAACGTTA	TGATGAAATC	GTCCAAGAAA	TGACACACAT	CCTTGATAAC	ATT---GGTT	TC-----	-----	[549]
050_EF1a_Trichomonas_vaginalis	-----	-----	---TACAACA	AGGCTCGTTT	CGACGAAATC	ACAGCAGAAA	TGACACGAT	CCTTACAGGC	ATC---GGCT	AC-----	-----	[549]
051_EF1a_Hexamita_inflata	-----	-----	---TACTCTG	AAGCTCGTTA	CACCTGATA	AAGACTGAAA	TGCAAAAGAC	TTTCAAGCAA	ATC---GGTT	TC-----	-----	[498]
052_EF1a_Spiroplasma_vortens	-----	-----	---TACTCCA	AGGACCGTTA	CAACGAGATC	AAGACCGAGA	TGACCAAGAC	CCTCGTCCGC	ATC---GGCT	AC-----	-----	[498]
053_EF1a_Andalucia_incarcerata	-----	-----	---TACGGCC	AGGCTCGTTT	TGAGGAAATC	CGCAAGGAAG	TTTCCGCTTT	CCTCAAGAAG	GTC---GGAT	AC-----	-----	[498]
054_EF1a_Seculamonas_ecuadoriensis	-----	-----	---TACTCGC	AGGCCGCTA	CGACGAAATC	AAGGCTGAGG	TGTCTGCCTA	CCTGAAGAAG	GTC---GGCT	AC-----	-----	[549]
055_EF1a_Reclinomonas_americana	-----	-----	---TACTCTG	AGGCTCGCTA	CAATGAGATC	AAGGCTGAGG	TGTCCGCTA	CCTGAAGAAG	GTC---GGCT	AC-----	-----	[549]
056_EF1a_Histiona_roides	-----	-----	---TACGGCC	AAGACCGTTA	CGACGAAATC	AAGAAGGAAG	TGTCTGCCTA	CCTCAAGAAG	GTC---GGCT	AC-----	-----	[549]
057_EF1a_Jakoba_libera	-----	-----	---TACTCCG	AGGCTCGCTA	CACCTGAGATC	AAGAAGGAGG	TTGGTGCTTA	CCTGAAGAAG	GTC---GGCT	AC-----	-----	[549]
058_EF1a_Stachyamoeba_lipophora	-----	-----	---TACTCTC	AGGACCGTTA	TGAGGAGATC	AAGAAGGAAG	TGTCTGACTA	CTTGAAGAAG	GTT---GGCT	AC-----	-----	[549]
059_EF1a_Naegleria_gruberi	-----	-----	---TATGCTG	AAAAGAGATA	TGACGAAATC	AAGTCTGAAG	TTGGTAGATA	TTTGAAGTCT	TTG---GGTT	TC-----	-----	[483]
060_EF1a_Naegleria_andersoni	-----	-----	---TACAAGG	AAGACAGATA	CAATGAAATC	AAGTCTGAAG	TTGGTAGATA	TTTGAAGGTT	TTG---GGCT	TC-----	-----	[498]
061_EF1a_Acrasis_rosea	-----	-----	---TACAAGG	AAGACCGTTA	CAAGGAAATC	CAAAAGGAAG	TGCTGACTA	CCTAAGAAG	GTC---GGAT	AC-----	-----	[498]
062_EF1a_Sawyeria_marylandensis	-----	-----	---TACTCAC	AAGAAAGATA	TACTGAAATT	AAAAATGGTG	TATCAGATAT	CATTAAAAAA	TAC---GGTT	AT-----	-----	[549]
063_EF1a_Bodo_saltans	-----	-----	---TTCTCCC	AGGCCCGCTA	CGACGAGATC	GTCTCCAACG	TTGGCCAGTA	CCTCAAGAAG	GTC---GGTT	AC-----	-----	[489]
064_EF1a_Rhynchobodo_sp.	-----	-----	---TACTCCA	AGGACCGTTA	TGAGGAGATC	ATGAAGGAGG	TCTCTGCCTA	CCTCAAGAAG	GTC---GGAT	AC-----	-----	[483]
065_EF1a_Rhynchomonas_nasuta	-----	-----	---TTCTCCG	AGGAGCGCTA	CGAGGAGATC	AAGAAGGAGG	TCTCGGCTT	CCTCAAGAAG	GTC---GGCT	AC-----	-----	[546]
066_EF1a_Dimastigella_trypaniformis	-----	-----	---TACAGCC	AGGACCGCTA	CAACGAGATC	GTCAAGGAGG	TGCTCTGCTA	CCTGAAGAAG	GTC---GGCT	AC-----	-----	[504]

	560	570	580	590	600	610	620	630	640	650	660	
067_EF1a_Perkinsiella_amoebae	-----	-----	---TACCAGC	AGGGCCGCTT	CGACGAGATC	AAGAAGGAAG	TCCAGCTCTA	CCTCAAGAAG	GTC--	--GGCT	AC-----	[474]
068_EF1a_Leishmania_major	-----	-----	---TACGCGC	AGTCACGCTA	CGATGAGATC	AGCAAGGAGG	TGGGCGCGTA	CCTGAAGCGC	GTG--	--GGCT	AC-----	[549]
069_EF1a_Leishmania_infantum	-----	-----	---TACGCGC	AGTCGCGCTA	CGATGAGATC	AGCAAGGAGG	TGGGCGCGTA	CCTGAAGCGC	GTG--	--GGCT	AC-----	[549]
070_EF1a_Leishmania_braziliensis	-----	-----	---TACTCGC	AGGCGCGCTA	TGAGGAGATC	AGCAAGGAGG	TGGGACAGTA	CCTGAAGCGG	GTG--	--GGCT	AC-----	[549]
071_EF1a_Leishmania_tarentolae	-----	-----	---TACTCGC	AGTCGCGCTA	CGATGAGATC	AGCAAGGAGG	TGGGCGCGTA	CCTGAAGCGC	GTG--	--GGCT	AC-----	[549]
072_EF1a "L. tarentolae" FJ807257	-----	-----	---TACAGCC	AGGACCGGTA	CAACGAGATT	GTGAAGGAAG	TGGGCACGTA	CCTGAAGAAG	GTT--	--GGGT	AC-----	[474]
073_EF1a_Leptomonas_costaricensis	-----	-----	---TACTCGC	AGGCGCGCTA	CGAGGAGATC	AGCAAGGAGG	TGGGCGCGTA	CCTCAAGCGC	GTC--	--GGCT	AC-----	[297]
074_EF1a_Leptomonas_bifurcata	-----	-----	---TACTCCC	AGGCCCGCTA	CGAGGAGATC	AGCAAGGAGG	TGGGCGCGTA	CCTGAAGCGC	GTT--	--GGCT	AC-----	[309]
075_EF1a_Leptomonas_podlipaevi	-----	-----	---TACGCC	AGGCCCGCTA	CGAGGAGATC	ACCAAGGAGG	TGGTGCCTA	CCTGAAGCGC	GTC--	--GGCT	AC-----	[297]
076_EF1a_Herpetomonas_muscarum	-----	-----	---TACAGCC	AGGACCGCTA	CAACGAGATC	ATGAAGGAGG	TGAGCGCTA	CGTGAAGAAG	GTC--	--GGCT	AC-----	[462]
077_EF1a_Herpetomonas_pessoai	-----	-----	---TTCTCCG	AGGACCGTTA	CAACGAGATC	AAGAAGGAGG	TCAGCACCTA	CCTCAAGAAG	GTC--	--GGCT	AC-----	[330]
078_EF1a_Blastocirithidia_culicis	-----	-----	---TGGAGCC	AGGACCGCTA	CAACGAGATC	ACCAAGGAGG	TGGTGCCTA	CCTGAAGAAG	GTC--	--GGCT	AC-----	[462]
079_EF1a_Trypanosoma_brucei	-----	-----	---TACGGAC	AGGACCGGTA	TGACGAGATT	GTGAAGGAGG	TGTCCTGCTA	CATCAAGAAG	GTT--	--GGGT	AC-----	[549]
080_EF1a_T. brucei_equiperdum	-----	-----	---TACGGAC	AGGACCGGTA	TGACGAGATT	GTGAAGGAGG	TGTCCTGCTA	CATCAAGAAG	GTT--	--GGGT	AC-----	[471]
081_EF1a_Trypanosoma_avium	-----	-----	---TACGGCC	AGGACCGCTA	CGACGAGATC	GTGAAGGAGG	TGTCACCTA	CCTCAAGAAG	GTC--	--GGCT	AC-----	[495]
082_EF1a_Trypanosoma_cruzi	-----	-----	---TTCCGCC	AGGACCGCTA	CGATGAGATT	GTGAAGGAGG	TGTCGGCGTA	CCTGAAGAAG	GTT--	--GGGT	AC-----	[549]
083_EF1a_Trypanosoma_rangeli	-----	-----	---TACAGCC	AGGCCCGCTA	CGAGGAGATT	GTGAAGGAGG	TGTCGGCGTA	CCTGAAGAAG	GTC--	--GGTT	AC-----	[549]
084_EF1a_Ploetia_costatum	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAGATC	AAGAAGGAAG	TTTCGACCTA	CTTGAAGAAG	GTC--	--GGCT	AC-----	[546]
085_EF1a_Ploetia_oblonga_cDNA	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAGATC	AAGAAGGAAG	TTTCGACCTA	CTTGAAGAAG	GTC--	--GGCT	AC-----	[546]
086_EF1a_Ploetia_oblonga_gDNA	-----	-----	---TACAACA	AGGATCGTTA	TGAGGAGGTC	AAGAAGGAGG	TGTCGGCGTA	CCTCAAGAAG	GTC--	--GGCT	AC-----	[546]
087_EF1a_Entosiphon_sp.	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAGATC	AAGAAGGAGG	TCAGCGCGTA	CCTGAAGAAG	GTC--	--GGCT	AC-----	[546]
088_EF1a_Entosiphon_sulcatum	-----	-----	---TACTCCC	AGGCCCGTTA	CGAAGAGATC	AAGAAGGAGG	TCAGCGCGTA	CCTGAAGAAG	GTC--	--GGCT	AC-----	[474]
089_EF1a_Peranema_tricophorum	-----	-----	---TACAACA	AGACACGATA	TGAGGAAATC	AAGAAGGAAG	TCAGTGCCTA	CCTCAAGAAA	GTC--	--GGAT	AC-----	[546]
090_EF1a_P. tricophorum_FJ807243	-----	-----	---TACAACA	AGACACGATA	TGAAGAAATC	AAGAAGGAAG	TGTCGCCTA	CCTCAAGAAA	GTT--	--GGAT	AC-----	[540]
091_EF1a_Colacium_elongatum	-----	-----	---TACTCTC	AGGCTCGGTA	TGAGGAAATT	AAGAAGGAGG	TGTCGGCTA	CCTTAAGAAG	GTT--	--GGTT	AC-----	[546]
092_EF1a_Colacium_mucronatum	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAGATC	AAGAAGGAGG	TGTCGGGATA	CCTCAAGAAG	GTT--	--GGCT	AC-----	[492]
093_EF1a_Colacium_sp.	-----	-----	---TACTCCC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAGG	TTTCGGGATA	CTTGAAGAAG	GTT--	--GGCT	AC-----	[546]
094_EF1a_Colacium_vesiculosum	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAGATC	AAGAAGGAGG	TGTCGGGATA	CCTCAAGAAG	GTT--	--GGCT	AC-----	[504]
095_EF1a_Cryptoglena_pigra	-----	-----	---TACTCTC	AGGCCCGTTA	TGAGGAAATC	AAGAAGGAGG	TGTCGGGATA	CCTGAAGAAG	GTT--	--GGCT	AC-----	[546]
096_EF1a_Cryptoglena_xujai	-----	-----	---TACTCTC	AGGCCCGTTA	TGAGGAAATC	AAGAAGGAGG	TGTCGGGATA	CTTGAAGAAG	GTT--	--GGGT	AC-----	[546]
097_EF1a_Discoploastis_adunca	-----	-----	---TTCCGTC	AGGCTCGCTA	TGAGGAAATT	AAGAAGGAGG	TGTCGGGATA	CCTCAAGAAG	GTT--	--GGTT	AC-----	[546]
098_EF1a_Euglena_clara	-----	-----	---TACTCTC	AGAGCCGATA	TGAGGAGATC	AAGAAGGAGG	TTTCGGGATA	CCTCAAGAAG	GTT--	--GGTT	AC-----	[426]
099_EF1a_Euglena_deses	-----	-----	---TATTAC	AGGCCCGATA	TGAGGAAATC	AAGAAGAAG	TGTCAGGATA	TCTGAAGAAG	GTG--	--GGAT	AC-----	[546]
100_EF1a_Euglena_exilis	-----	-----	---TATTAC	AGGCACGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCATGTA	TTTGAAGAAG	GTG--	--GGCT	AC-----	[426]
101_EF1a_Euglena_geniculata	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAGATC	AAGAAGGAAG	TTTCGACCTA	CTTGAAGAAG	GTC--	--GGCT	AC-----	[546]
102_EF1a_Euglena_gracilis	-----	-----	---TACTCCC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAGG	TGTCGGGATA	CTTGAAGAAG	GTT--	--GGCT	AC-----	[546]
103_EF1a_E. gracilis_X16890	-----	-----	---TACTCTC	AGGCCCGTTA	TGAAGAAATC	AAGAAGGAGG	TTTCGGGATA	TTTGAAGAAG	GTG--	--GGCT	AC-----	[549]
104_EF1a_E. gracilis_FJ807238	-----	-----	---TACTCCC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAGG	TGTCGGGATA	CTTGAAGAAG	GTT--	--GGCT	AC-----	[549]
105_EF1a_Euglena_granulata	-----	-----	---TACTCCC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAGG	TTTCACCTA	CCTCAAGAAG	GTG--	--GGTT	AC-----	[546]
106_EF1a_Euglena_hiemalis	-----	-----	---TACTCCC	AGTCCCGATA	TGAGGAGATC	AAGAAGGAAG	TCTCTGGTA	CTTGAAGAAG	GTT--	--GGCT	AC-----	[546]
107_EF1a_Euglena_longa	-----	-----	---ACTC	AAGGCTCAT	TGAAGAGTTC	AAGAAGGAAG	TTTGTGGATA	CTTGAAGAAA	GTT--	--GGTA	AT-----	[63]
108_EF1a_Euglena_stellata	-----	-----	---TATTCCC	AGTCTCGTTA	TGAAGAGATC	AAGAAGGAAG	TTTCTGGGTA	CCTCAAGAAG	GTT--	--GGTT	AC-----	[426]
109_EF1a_Euglena_viridis	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAGATC	AAGAAGGAAG	TTTCGGGATA	TTTGAAGAAG	GTT--	--GGTT	AC-----	[504]
110_EF1a_Eutreptia_pertyi	-----	-----	---TACTCCC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAGG	TTTCACCTA	CCTCAAGAAG	GTG--	--GGTT	AC-----	[546]
111_EF1a_Eutreptiella_pomquetensis	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAGATC	AAGAAGGAAG	TTTCGACCTA	CTTGAAGAAG	GTC--	--GGCT	AC-----	[156]
112_EF1a_Lepocinclis_acus_var._major	-----	-----	---TACTCTC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCGGCTA	CTTGAAGAAG	GTC--	--GGTT	AC-----	[546]
113_EF1a_Lepocinclis_ovum	-----	-----	---TATTAC	AGGCTCGTTA	TGAAGAAATC	AAGAAGGAGG	TGTCGGCTA	CTTGAAGAAG	GTT--	--GGAT	AC-----	[426]
114_EF1a_Lepocinclis_spyrogyroides_clone_38	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
115_EF1a_L. spyrogyroides_clones_1_&_4	-----	-----	---TACTCTC	AGGCCCGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCACCAA	TCTGAAGAAG	GTG--	--GGCT	AC-----	[546]
116_EF1a_Lepocinclis_steinii	-----	-----	---TACAATG	AGGCTCGGTA	TGAGGAAATC	AAGAAGGAGG	TGTCGGGATA	CTTGAAGAAG	GTT--	--GGAT	AC-----	[426]
117_EF1a_Lepocinclis_triapteris	-----	-----	---TATTCCC	AGGCACGTTA	TGAGGAGATC	AAGAAGGAGG	TGTCGGGATA	CCTGAAGAAG	GTC--	--GGAT	AC-----	[546]
118_EF1a_Monomorphina_aenigmatica	-----	-----	---TACTCCC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAGG	TGTCGGGATA	TTTGAAGAAG	GCG--	--GGTT	AC-----	[546]
119_EF1a_Monomorphina_pyrum	-----	-----	---TACTCTC	AGGCCCGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCACCTA	CTGAAGAAG	GTG--	--GGCT	AC-----	[546]
120_EF1a_Monomorphina_sp.	-----	-----	---TACTCTC	AGGCCCGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCGGGATA	CCTTAAGAAG	GTT--	--GGTT	AC-----	[546]
121_EF1a_Phacus_orbicularis	-----	-----	---TACTCCC	AGGCTCGCTA	TGAGGAAATC	AAGAAGGAGG	TGTCGCTA	CCTGAAGAAG	GTT--	--GGCT	AC-----	[504]
122_EF1a_Phacus_similis	-----	-----	---TACTCTC	AGTCCCGTTA	TGAGGAGATC	AAGAAGGAGG	TGTCGGGATA	CTTGAAGAAG	GTT--	--GGCT	AC-----	[546]
123_EF1a_Strombomonas_acuminata	-----	-----	---TATTCCC	AGGCACGATA	TGAGGAGATC	AAGAAGGAGG	TGTCGGGATA	CTTGAAGAAG	GTT--	--GGCT	AC-----	[549]
124_EF1a_Strombomonas_borysteniensis	-----	-----	---TACTCCC	AGGCCCGTTA	TGAGGAAATC	AAGAAGGAAG	TTTCTGGGATA	CCTCAAGAAG	GTT--	--GGTT	AC-----	[426]
125_EF1a_Strombomonas_napiformis	-----	-----	---TACTCCC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCGGGATA	CTTGAAGAAG	GTG--	--GGTT	AC-----	[546]
126_EF1a_Strombomonas_ovalis	-----	-----	---TACTCTC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCGGGATA	CCTCAAGAAG	GTT--	--GGTT	AC-----	[546]
127_EF1a_Strombomonas_triquetra	-----	-----	---TACTCCC	AGGCACGTTA	TGAGGAAATC	AAGAAGGAGG	TGTCGGGATA	CCTCAAGAAG	GTT--	--GGTT	AC-----	[546]
128_EF1a_Trachelomonas_conradi	-----	-----	---TACTCCC	AGGCCCGCTA	CGAAGAAATC	AAGAAGGAAG	TGTCAGGATA	CTTGAAGAAG	GTT--	--GGCT	AC-----	[546]
129_EF1a_Trachelomonas_echinata	-----	-----	---TACTCAC	AGGCTCGTTA	TGAAGAAATC	AAGAAGGAGG	TTTCAGGCTA	CCTGAAGAAG	GTT--	--GGAT	AC-----	[546]
130_EF1a_Trachelomonas_ilefvelei	-----	-----	---TACTCAC	AGGCTCGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCCTCTA	CCTGAAGAAG	GTT--	--GGGT	AC-----	[546]
131_EF1a_Trachelomonas_planconica	-----	-----	---TACTCAC	AGAGTCGTTA	TGAGGAAATC	AAGAAGGAAG	TGTCACCTA	CCTGAAGAAG	GTG--	--GGTT	AC-----	[426]
132_EF1a_Trachelomonas_similis	-----	-----	---TATTACG	AGGCTCGCTA	TGAGGAAATC	AAGAAGGAGG	TGTCGGGATA	CCTCAAGAAG	GTT--	--GGTT	AC-----	[426]

	560	570	580	590	600	610	620	630	640	650	660	
133_EFL1a_Trachelomonas_volvocina	-----	-----	---TACTCAC	AGGCCCGTTA	CGAAGAAATC	AAGAAGGAAG	TGTCCATGTA	CTTGAAGAAG	GTT---GGTT	AC-----	-----	[546]
134_EFLenvironmental_sample_EU810232	-----	-----	---TACGCA	AGGAGCGCTT	TGAGGAGATC	AAGGACGAGA	TGACCAAGAT	GTTGACCAAG	GTC---GGCT	TGAAGAAAGG	CTTTGTC---	[585]
135_EFLCreolimax_fragrantissima	-----	-----	---TACGGTA	TCGAGCGTTA	CAGTGAAGAT	AAGGACGAGG	TTGTTCATAT	GTTGACCCCG	GTC---GGAT	GGAAGAAGGA	CTTTGTT---	[513]
136_EFLSpaeroforma_arctica	-----	-----	---TACGGTA	TCGAGAGATA	CAGTGAAGAT	AAGGATGAGG	TGTTACATAT	GTTGACCCCG	GTT---GGAT	GGAAGAAGGA	TTTCGTT---	[585]
137_EFLMonosiga_brevicollis	-----	-----	---TACGGC	AGGCCCGCTT	CAAGGAGGTC	CGCGACGAGA	TGGTGACAT	GTTGACCAAT	GTC---GGCT	GGAAGAAGG	CTTCGTT---	[585]
138_EFLXenopus_laevis	-----	-----	---TACGGCA	AGGCCCGCTA	CGATGAGATC	AAGGACGAGA	TGATCCATAT	GTTTACCCCG	GTC---GGCT	GGAAGAAGGA	CTTTGTC---	[585]
139_EFLSpizellomyces_punctatus	CC-----	-----	---TACAAGA	AGGAGCGTTA	CGATGAAATC	GCCAACGAGA	TGCGCAACAT	GTTGGTCCGT	ACC---GGCT	GGAAGAAGGA	GTTTCGTT---	[588]
140_EFLGaertneriomyces_semiglobiferous	CC-----	-----	---TACAAGA	AGGAGCGTTA	CGATGAAATC	GCCAACGAGA	TGCGTCCAT	GTTGGTCCGT	ACC---GGCT	GGAAGAAGGA	CTTTGTC---	[291]
141_EFLAllomyces_macrogynus	CC-----	-----	---TACAAGA	AGGAGCGCTA	CGACGAAAT	GCCGCTGAGA	TGAAGGACAT	GCTCGTCAAG	ACG---GGCT	GGAAGAAGGA	CTTCGTC---	[588]
142_EFLBlastocladiella_emersonii	CC-----	-----	---TACAAGA	AGGAGCGCTA	CGACGAGATC	TCGGCCGAGA	TGAAGGACAT	GCTCATCAAG	ACC---GGCT	GGAAGAAGGA	CTTCGTC---	[588]
143_EFLCatenophylctis_sp.	CC-----	-----	---TACAAGA	AGGAGCGTTA	CGACGAAATC	TCGGCTGAGA	TGAAGGACAT	GCTCGTCAAG	ACC---GGT	GGAAGAAGGA	GTTTGT---	[291]
144_EFLCoelomomyces_stegomyiae	CT-----	-----	---TACAAC	AGGCGCGTTA	CGATGAAAT	GCTGCTGAAA	TGGCCGATAT	GTTGGTTAAA	ACT---GGAT	GGAAGAAGGA	ATTTGTT---	[291]
145_EFLEntomophthora_muscae	-----	-----	---TACAAG	AGGATCGTTA	CAACGAGATC	CGTGATGAGA	TGCGCTCCAT	GTTGATCAAG	GTC---GGAT	GGAAGAAGGA	GTTTCGTT---	[288]
146_EFLBasidiobolus_ranarum	-----	-----	---TACAAG	AAGACCGTTA	TGTTGAAATC	CGTGATGAGA	TGCAAAACAT	GTTGATCAAG	GTC---GGAT	GGAAGAAGGA	WTTTCGTT---	[288]
147_EFLConidiobolus_coronatus	-----	-----	---TACAAG	AGGCCCGTTA	CGTCAAAATC	CGTGACGAAA	TGATCAATAT	GTTAACCAAG	GTC---GGAT	GGAAGCCCGA	TTTCATC---	[288]
148_EFLTetraselmis_tetratele	-----	-----	---CAACAAG	AGTGCAGCTA	CGTCGAGATC	CGCGATGAGA	TGATCAACAC	CCTCGTCAAG	GTT---GGCT	GGAAGAAGGA	CTTCGTC---	[534]
149_EFLMesostigma_viride	-----	-----	---TACAAGA	AGGAGCGCTA	CAGTGAAT	TCCGATGAGA	TGAAGAACAT	GTTTGTAGG	GTG---GGCT	GGAAGAAGGA	GTTTCGTT---	[588]
150_EFLOstreococcus_tauri	-----	-----	---TACAAG	AAGAGCGTTA	CGAAGAAAT	CGCAACGAAA	TGAAGAACAT	GTTGATCAAG	GTC---GGCT	GGAAGAAGGA	CTACGTC---	[585]
151_EFLNephroselmis_olivacea	-----	-----	---TTCAAG	AGTCTCGCTA	CAGTGAAGAT	CGCGATGAGA	TGCGCAACAT	GCTCTGAAAG	GTG---GGCT	GGAAGAAGGA	GTTTGT---	[561]
152_EFLMicromonas_pusilla	-----	-----	---TACAAG	AGGAGCGCTA	CGAGGAGATC	CGCAACGAGA	TGAAGAACAT	GCTCATGAAG	GTG---GGCT	GGAAGAAGGA	CTACGTC---	[534]
153_EFLParachlorella_kessleri	-----	-----	---TACAAG	AAGAGCGCTA	CAAGGAGAT	GCCGACGAGA	TGAAGCACAT	GCTTGTGCGC	GTG---GGCT	GGAAGGCGGA	CTTTGTG---	[543]
154_EFLHelicosporidium_sp.	-----	-----	---TACTCCA	AGGAGCGCTA	CGACGAGATC	TCCAAGGAGA	TGCGCTTCAT	GCTCGTGAAG	GTG---GGCT	GGAAGGCGGA	CTTCGTT---	[585]
155_EFLGonium_pectorale	-----	-----	---TACAAG	AGGCCCGCTA	CAAGGAGAT	GCCGAGGAGA	TGAAGCACAT	GCTCGTCCGT	GTG---GGCT	GGAAGGACAC	CTTCGTC---	[585]
156_EFLChlamydomonas_reinhardtii	-----	-----	---TACAAGA	AGGAGCGCTA	CGATGAGATC	GCCAACGAGA	TGCGCCACAT	GCTGGTGC	GTG---GGCT	GGAAGGACGA	CTTCGTC---	[585]
157_EFLScenedesmus_obliquus	-----	-----	---TACAGG	AGGAGCGCTA	TGTTGAGATC	CGCGACGAGA	TGCGCCACAT	GCTTGC	GTG---GGCT	GGAAGCCTGA	CTTCATT---	[585]
158_EFL"Oryza_sativa"	-----	-----	---TACAAG	AGGAGCGCTA	CGTTGAGATC	CGTGATGAGA	TGAAGCACAT	GCTTACGAG	GTC---GGCT	GGAAGCCCGT	CTTTGTG---	[585]
159_EFLChlorococcum_sp.	-----	-----	---TACAAG	AGGAGCGCTA	CAACGAGATC	AAGGATGAGA	TGAAGCACAT	GCTTATCCGC	GTC---GGCT	GGAAGAAGGA	GTACCTG---	[534]
160_EFLOchlochaete_hystrix	-----	-----	---TACAAG	AGGAGCGCTA	CAAGGAGAT	TCCGACGAGA	TGAAGCACAT	GCTTGCCCGC	GTT---GGCT	GGAAGCCGGA	GTTTCGT---	[543]
161_EFLUlvaintestinalis	-----	-----	---TACAAG	AGGCTCGCTA	CAGTGAAGAT	AAGGATGAGA	TGAAGCACAT	GCTCATGAAG	GTT---GGAT	GGAAGAAGGA	GTTTGT---	[534]
162_EFLUrospora_sp.	-----	-----	---TACAAG	AGGCGCGCTA	CAACGAGATC	AAGGACGAGA	TGAAGCACAT	GCTACCAAG	GTG---GGCT	GGAAGAAGGA	GTTTGT---	[387]
163_EFLGymnochlorella_stellata_ppc	-----	-----	---TATAGTA	AAGCAAGATA	TGATGAGATT	CTGAAGAGAT	TTAAGTCCAC	ACTAGCAAG	GTT---GGAT	GGAAGAAGGA	CTTCATC---	[621]
164_EFLLotharella_amoebiformis_ppc	-----	-----	---TACTCGG	AGGATCGCTA	CGAAGAGACT	CTAAGAGGCT	TAAAGGCGCT	CCTCAAAAG	GTG---GGAT	GGAAGAAGGA	CTTCATC---	[366]
165_EFLChlorarachnion_reptans_ppc	-----	-----	---TACTCTG	AGAAACGATT	CAATGAGATC	TTGAAGGAAC	TGAAGTCTG	CCTCGCAAAA	GTT---GGAT	GGAAGAAGGA	CTTCCTA---	[393]
166_EFLChondrus_crispus	-----	-----	---TACGCCA	AGCAGCGATT	CAGTGAAGAT	AAGAATGAGG	TTACGAGCAT	GCTTGTCAAG	GTT---GGGT	GGAAGAAGGA	CTTGTG---	[585]
167_EFLGracilaria_changii	-----	-----	---TATGGTA	AGGACGATT	CGACGAGATT	CGCGCCGAGG	TCACACTACT	GCTCATGAAA	GTT---GGAT	GGAAGAAGGA	TTTCGTT---	[585]
168_EFLCryptomonas_ovata	-----	-----	---TACAAG	AGGACCGTTA	CGAGGAGATC	CGCAACGAGA	TGCGCAACAT	GCTTATCCGT	GTG---GGGT	GGAAGAAGGA	TTTCGTC---	[540]
169_EFLGuillardia_theta	-----	-----	---TACAAG	AGGAGCGCTA	CAACGAGATC	CGCGATGAGA	TGAGGAACAT	GCTCATCCGT	GTT---GGAT	GGAAGAAGGA	GTTTCGTC---	[582]
170_EFLRhodomonas_salina	-----	-----	---TACAAG	AGGCTCGCTA	CAACGAGATC	GCGCAGGAGA	TGCGCAACAT	GCTTATCCGC	GTG---GGGT	GGAAGAAGGA	CTTCGTC---	[582]
171_EFLPavlova_lutheri	CG-----	-----	---TACTCCA	AGGCGCGGTA	CGACGAGGTC	GCGCGCGAGA	TGAAGTCGAT	GCTCATCAAG	GTC---GGCT	GGAAGAAGGA	GTTTCGTC---	[585]
172_EFLChrysochromulina_sp.	CG-----	-----	---TACGCCA	AGGCCCGCTA	CGACGAGGTC	GCGCGCGAGA	TGAAGAACAT	GCTCATCAAG	GTC---GGCT	GGAAGAAGGA	CTTCGTC---	[405]
173_EFLIsochrysis_galbana	-----	-----	---TACAAGA	AGTCCCGCTA	TGACGAGGTC	GCGCGCGAGA	TGAAGAACAT	GCTTATCAAG	GTC---GGCT	GGAAGAAGGA	CTTTGTT---	[585]
174_EFLEmiliana_huxleyi	CG-----	-----	---TACTCCA	AGGCGCGCTA	CGACGAGGTC	GCGCGCGAGA	TGAAGAACAT	GCTTCTCAAG	GTC---GGCT	GGAAGAAGGA	CTTTGTT---	[585]
175_EFLAchnanthes_kuwaitensis	-----	-----	---TGGGACG	AGGGTCGTTT	CAACGAGATC	AAGGAAGAGT	TCGTGAAGAT	GCTTCAGCTC	ATC---GGCT	TCAAG-----	-----	[504]
176_EFLAsterionella_glacialis	-----	-----	---TGGGACC	AAGGACGTTT	CAACGAGATC	AAGGACGAAT	TACCAAGAT	GCTTCAATG	ATC---GGCT	TCAAG-----	-----	[504]
177_EFLDetonula_confervacea	-----	-----	---TGGTCCG	AGGAACGATT	CAACGAGATC	AAGGAGGAGT	TCCAGAAGAT	GCTTCAATTG	ATC---GGCT	TCAAG-----	-----	[504]
178_EFLThalassiosira_pseudonana	-----	-----	---TGGGACG	AGGGACGTTT	CAACGAGATT	AAGGACGAGT	TCCAGAAGAT	GCTTCAATG	ATT---GGGT	TCAAG-----	-----	[567]
179_EFLPerkinsus_marinus	-----	-----	---TACAAG	AATCTCGTTA	TATTGAGATT	GCGCAGGAGA	TGAAGAATGT	TCTTGCGAGA	GTT---GGTT	GGAAGAAGGA	CTTCGTT---	[594]
180_EFLOxyrrhis_marina	-----	-----	---TACAAG	AGGCTCGRTA	CGATGAGATC	GCCAACGAGA	TGAGGAACAT	GCTTGCCAAG	GTC---GGGT	GGAAGAAGGA	TTTCCTT---	[591]
181_EFLKarlodinium_micrum	-----	-----	---TACAAC	AGGCTCGATA	TGATGAGATT	GCAAAATGAGA	TGAGGAGCAT	GTTGGTGAAG	GTT---GGTT	GGAAGAAGGA	TTTGTGA---	[591]
182_EFLAmphidinium_carterae	-----	-----	---TACAAGA	AGGACCGATA	CGATGAGATT	TCCAACGAGA	TGAAGAGCAT	GTTGGTGAAG	GTG---GGTT	GGAAGAAGGA	CTTTGTG---	[591]
183_EFLHeterocapsa_triquetra	-----	-----	---TACAAG	AGTCCGCGTA	CGATGAGATT	GCGAACGAGA	TGAAGAGCAT	GTTTGTGAAG	GTT---GGCT	GGAAGAAGGA	CTTCGTC---	[591]
184_EFLPlanoglabratella_operculis	-----	-----	---TATGGAG	AAGGACGATA	CAATGAGATC	AAGAGTGAGG	TCGATAAAAT	GCTTACCAAG	ATC---GGAT	ATAAG-----	-----	[411]
185_EFLReticulomyxa_filosa	-----	-----	---TATGCTC	AAGATCGTTT	TGATGAAATC	AAATCGGAAG	TCGAGAAAAT	GTTAACCAAG	ATT---GGAT	ACAAA-----	-----	[501]
186_EFLLotharella_amoebiformis	-----	-----	---TTGACC	AAAAGCGCTA	CAAGGAGATC	AAGTCCAACA	TGCTTAAAGT	GCTTAAACAG	GCC---GGAT	GGAAGACCCA	GGAAGAGCCT	[633]
187_EFLGymnochlorella_stellata	-----	-----	---TACGAT	AAGTTCGTTA	CAAGGAGATC	AAAAAGAATA	TGCTTTCTAT	GCTTAAAGCA	GCT---GGAT	GGAAGATCAA	GGAAGAGCCA	[549]
188_EFLChlorarachnion_reptans	-----	-----	---TGGGACC	AAGGTTCGTTA	CAAGGAAATC	AGGAAGAATA	TGCTTTCTAT	GCTTAAAGCA	TCC---GGTT	GGAAGATCAA	GGAAGAGTCC	[516]
189_EFLDiplonema_ambulator	-----	-----	---TACAAG	AGGAGAGGTA	CAACGAGATC	AAGGAGGAGA	TGAAGAGCAT	GTTGGTGAAG	ACG---GGGT	GGAAGAAGGA	GTTTCGTC---	[186]
190_EFLDiplonema_papillatum	-----	-----	---TACAAG	AGGAGCGATA	CAACGAGATC	AAGGAGGAGA	TGAAGAGCAT	GATGATCAAG	ACT---GGCT	GGAAGAAGGA	CTTCATT---	[558]
191_EFLRhynchopus_euleeides	-----	-----	---TACAAG	AGGAGCGCTT	CAACGAGATC	AAGGAGGAGA	TGAAGAGCAT	GCTCATGAAG	ACG---GGGT	GGAAGAAGGA	CTTCGTC---	[198]
192_EFLPetalomonas_cantuscygni	-----	-----	---TACTCCA	AGGAGCGTTA	CGACGAGGTC	TCTCTGAGA	TGCGCTGCAC	CCTCGTCCG	GTC---GGCT	GGAAGAAGGA	CTTCGTC---	[429]
193_EFLTrypanoplasma_borrel	GCCAACCC--	-----	---TACAGYC	AGGCCCGYTA	CAATGAGGTY	GCKAGCGAGA	TGAAGCGCAT	TCTGCTGCAG	GAG---GGAT	GGAAGAAGGA	CCAGATC---	[603]
194_EFLNeobodo_salien	AGACG----	-----	---TACAGCG	AGAAAGCGTT	CATCGAGATC	TCGACCGAGG	TCAAGCGCAT	TCTGCTGCAG	GAG---GGCT	GGAAGAAGGA	TCAGATC---	[549]

	670	680	690	700	710	720	730	740	750	760	770	
001_aEF1a_Desulfurococcus_mobilis						-GAT		GT	CTCGAAGATA	CCCTTCATAC	CTATATCGGC	[587]
002_aEF1a_Sulfolobus_solfataricus						-AAT		AC	AAATAAGGTT	AGATTTGTAC	CAGTAGTTGC	[581]
003_aEF1a_Thermoplasma_acidophilum									-AaGgAtAta	tCaTtCgTgc	cCaTaAgCgG	[563]
004_aEF1a_Halobacterium_sp.						-AAC		CC	CGACGACGCG	AAGTTCATCG	CGACCTCCGG	[554]
005_aEF1a_Methanocaldococcus_jannaschii						-AAC		CC	AGACCAAAAT	GACTTCATCC	CAACAGCTTC	[569]
006_EF1a_Acanthamoeba_culbertsoni						-AAC		CC	CAAGAACATC	CCGTTCTGTC	CCATCTCCGG	[551]
007_EF1a_Neoparamoeba_branchipila						-AAC		CC	CGCCAAGATC	CCCTTCGTCC	CCATCTCCGG	[536]
008_EF1a_Planoprotostelium_aurantium						-AAC		CC	CGAGAAGGTT	CCCTTCGTCC	CCATCTCCGG	[533]
009_EF1a_Physarum_polycephalum						-AAC		CC	CGAGAAGATC	GCCTTCGTCC	CCATCTCCGG	[533]
010_EF1a_Amoebidium_parasiticum						-AAC		CC	CAAGGCCGTT	CCCTTCGTGC	CCATCTCCGG	[548]
011_EF1a_Capsaspora_owczaraki						-GAC		CC	CAAGACCGTC	GCTTTCGTCC	CCATCTCTGG	[578]
012_EF1a_Monosiga_ovata						-AAC		CC	CGACACTGTT	GCCTTCGTCC	CCATCTCTGG	[557]
013_EF1a_Salpingoeca_amphoridium						-AAC		CC	CGACACTGTT	GCCTTTGTCC	CCATCTCGGG	[551]
014_EF1a_Ceanorhabditis_elegans						-AAT		CC	AAAGGCTGTT	CCATTCTGTC	CAATCTCTGG	[584]
015_EF1a_Drosophila_melanogaster						-AAC		CC	AGCCGCCGTT	GCCTTCGTGC	CAATTTCCGG	[584]
016_EF1a_Ornithorhynchus_anatinus						-AAC		CC	TGCCACAGTC	CCCTTTGTGC	CCATCTCCGG	[584]
017_EF1a_Xenopus_laevis						-AAC		CC	AGCGACCGTT	CCATTTGTGC	CAATATCTGG	[584]
018_EF1a_Saccharomyces_cerevisiae						-AAC		CC	AAAGACTGTT	CCATTCTGTC	CAATCTCTGG	[578]
019_EF1a_Neurospora_crassa						-AAC		CC	CGCTGGTGTG	GCTTTCGTCC	CCATCTCCGG	[581]
020_EF1a_Cryptococcus_neoformans						-AAC		CC	CAAGGCTGTC	CCCTTCGTCC	CCATCTCTGG	[578]
021_EF1a_Schizophyllum commune						-AAC		CC	GAAGACCGTC	GCCTTCGTCC	CGATCTCCGG	[578]
022_EF1a_Monoblepharis_insignis						-AAC		CC	CAAGACGGTT	GCCTTCGTGC	CTATCTCGGG	[464]
023_EF1a_Rhizophlyctis_harderi						-AAC		CC	AAAGGCTGTT	CCTTTCATCC	CAATCTCTGG	[464]
024_EF1a_Harpochytrium_sp.						-AAC		CC	CGCCGCCGTT	CCGTTCTGTC	CCATCTCCGG	[416]
025_EF1a_Absidia_glauca						-AAC		CC	CAAGTCTGTC	CCCTTCGTCC	CCATCTCTGG	[578]
026_EF1a_Smittium_simulii						-AAT		CC	TAAAGCTGTT	GCCTTTGTTC	CTATCTCTGG	[551]
027_EF1a_Chara_australis						-AAC		CC	TGACAAGATA	AACCTCGTGC	CCATTTCTGG	[533]
028_EF1a_Physcomitrella_patens						-AAC		CC	GGAGAAGGTT	CCGTTCTGTC	CGATCTCAGG	[584]
029_EF1a_Acetabularia_acetabulum						-AAC		CC	CGAAAAAATT	AACCTTTGTGC	CGATTTCCGG	[533]
030_EF1a_Chaetomorpha_coliformis						-AAG		CC	TGACAACGTT	CCTTTCATCC	CCATCTCCGG	[383]
031_EF1a_Arabidopsis_thaliana						-AAC		CC	TGACAAAATC	CCATTTGTGC	CAATCTCTGG	[584]
032_EF1a_Oryza_sativa						-AAC		CC	TGACAAGATT	CCCTTTGTTC	CCATCTCTGG	[584]
033_EF1a_Cyanidioschyzon_merolae						-AAC		CC	GGACAAGGTT	CCCAAGGTAC	CAATCTCCGG	[584]
034_EF1a_Porphyrha_yezoensis						-AAC		CC	CGCCAAGGTT	CCCAAGGTGC	CGACGTCCGG	[584]
035_EF1a_Plasmodium_falciparum						-CAA		GC	TGATAAAGTT	GATTTTATTC	CAATTTCTGG	[578]
036_EF1a_Cryptosporidium_parvum						-AAC		AC	CGAGAAGATC	CCATTCTGTT	CCATTTCTGG	[572]
037_EF1a_Tetrahymena_pyriformis						-AAG		CC	CGACACCATC	CCCTTCATCC	CCATCTCTGG	[587]
038_EF1a_Stentor_coeruleus						-AAC		CC	AGACAAGATT	CCGTTCTATCC	CTATCTCAGG	[539]
039_EF1a_Phytophthora_parasitica						-AAG		CC	CGCCAAGATC	CCGTTCTGTC	CCATCTCTGG	[584]
040_EF1a_Mallomonas_rasilis						-AAG		CC	TGCCAAGATC	AACCTCGTTC	CAATTTCTGG	[545]
041_EF1a_Phaeodactylum_tricornutum						-AAG		CC	GATGAAGATT	CCATTCTGTC	CCATCTCCGG	[584]
042_EF1a_Thalassiosira_pseudonana						-AAG		AT	TGACAAAGTT	AAATTCTATCC	CCATCTCTGC	[584]
043_EF1a_Malawimonas_jakobiformis						-AAC		TT	GGAGAAGGTT	CAGTTCGTGC	CCATCTCCGG	[515]
044_EF1a_Malawimonas_californiana						-AAC		CC	CGAGAAGATC	CCCTTCGTGC	CCATCTCCGG	[584]
045_EF1a_Trimastix_pyriformis						-AAC		CT	CGAGAAGGTT	CAGGCCGTCC	CCATCTCTGG	[542]
046_EF1a_Dinenympha_exilis						-AAT		CC	GGACAAAATC	CTGATGATCC	CAATTTCAAG	[524]
047_EF1a_Monocercomonoides_sp.						-AAC		CC	AGATAAAAATC	CCAGTCATTC	CAATCTCTGG	[584]
048_EF1a_Streblomastix_strix						-AAC		CC	AGACAAAATC	CAAATGATCC	CAATCAGCGG	[533]
049_EF1a_Histomonas_meleagridis						-AAG		CC	AGAACAAATC	AAGTTCGTCC	CAATCTCTGG	[584]
050_EF1a_Trichomonas_vaginalis						-AAG		CC	AGAAATGTTC	CGCTTCGTCC	CAATCTCCGG	[584]
051_EF1a_Hexamita_inflata						-AAG		CACTG	GGAAGAGTTC	GACTTTGTCC	CACCTCTCCG	[536]
052_EF1a_Spironucleus_vortens						-AAG		AAGGC	CCCAGAGTTC	AACATCATCC	CGACCTCTGG	[536]
053_EF1a_Andalucia_incarcerata						-AAC		CC	TGCTAAGGTT	CCATTCTGTC	CTGTCTCTGG	[533]
054_EF1a_Seculamonas_ecuadoriensis						-AAC		CC	GGAGAAGGTC	AACTTTGTGC	CGATCTCGGG	[584]
055_EF1a_Reclinomonas_americana						-AAC		CC	GGACAAGGTC	AACCTCATCC	CCATCTCCGG	[584]
056_EF1a_Histiona_aroides						-AAC		CC	CGACAAGGTC	AACCTCATCC	CAATCTCCGG	[584]
057_EF1a_Jakoba_libera						-AAC		CC	CGAGCAGGTC	AACCTCGTGC	CCATCTCCGG	[584]
058_EF1a_Stachyamoeba_lipophora						-AAG		CA	GGAGATT---	CCTTTCATCC	CCATCTCCGG	[581]
059_EF1a_Naegleria_gruberi						-CAA	---ACTGAAG	AAGGTAAGCC	AACTGTGTG	CAATTCTGTC	CAATTTCTGG	[533]
060_EF1a_Naegleria_andersoni						-AAC	GTTGATGCTG	CTGATAAGCC	AACTTGGTC	CAATTCTGTC	CAATTTCTGG	[551]
061_EF1a_Acrasis_rosea						-AAC		CC	AAAGAATGTT	CCATTCTGTC	CAATTTCAAG	[533]
062_EF1a_Sawyeria_marylandensis						-AAA		CC	AGAAAGTGTT	CCATTTGTTC	CAATTTCAAG	[584]
063_EF1a_Bodo_saltans						-AAC		AT	CGAGAAGGTC	CGCTTCGTCC	CCATCTCCGG	[524]
064_EF1a_Rhynchobodo_sp.						-AAC		GT	CGAGAAGGTC	AAGTTCGTCC	CCATCTCCGG	[518]
065_EF1a_Rhynchomonas_nasuta						-AAC		GT	CGAGAAGATC	CGCTTCGTGC	CCATCTCCGG	[581]
066_EF1a_Dimastigella_trypaniformis						-AAC		GT	CGAGAAGGTC	AAGTTCATCC	CCATCTCCGG	[539]

	670	680	690	700	710	720	730	740	750	760	770	
067_EF1a_Perkinsiella_amoebae	-----	-----	-----	-----	-----	-AAC	-----	-----AT	CGAGAAGGTC	CAGTTCGTGC	CAATCTCCGG	[509]
068_EF1a_Leishmania_major	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTG	CGCTTCATCC	CGATCTCGGG	[584]
069_EF1a_Leishmania_infantum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTG	CGCTTCATCC	CGATCTCGGG	[584]
070_EF1a_Leishmania_braziliensis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTG	CGCTTCATCC	CGATCTCGGG	[584]
071_EF1a_Leishmania_tarentolae	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTG	CGCTTCATCC	CGATCTCGGG	[584]
072_EF1a "L. tarentolae" FJ807257	-----	-----	-----	-----	-----	-AAT	-----	-----AT	TGACAAAGGTG	CAGTTCGTGC	CAATCTCGGG	[509]
073_EF1a_Leptomonas_costaricensis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTG	CGCTTCATCC	CGATCTCGGG	[332]
074_EF1a_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTG	CGCTTCATCC	CGATCTCGGG	[344]
075_EF1a_Leptomonas_podlipaevi	-----	-----	-----	-----	-----	-AAC	-----	-----CT	GGAGAAGGTG	CGTTTCATCC	CAATCTCTGG	[332]
076_EF1a_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTG	AAGTTCATCC	CGATCTCTGG	[497]
077_EF1a_Herpetomonas_pestoi	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTG	AAGTTCATCC	CGATCTCTGG	[365]
078_EF1a_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-AAC	-----	-----GT	GGAGAAGGTG	CGCTTCATCC	CGATCTCCGG	[497]
079_EF1a_Trypanosoma_brucei	-----	-----	-----	-----	-----	-AAC	-----	-----GT	GGAGAAGGTG	CGCTTCGTCC	CCATCTCCGG	[584]
080_EF1a_T. brucei_equiperdum	-----	-----	-----	-----	-----	-AAC	-----	-----GT	GGAGAAGGTG	CGCTTCGTCC	CCATCTCCGG	[506]
081_EF1a_Trypanosoma_avium	-----	-----	-----	-----	-----	-AAC	-----	-----GT	GGAGAAGGTG	CGCTTCATCC	CGATCTCCGG	[530]
082_EF1a_Trypanosoma_cruzi	-----	-----	-----	-----	-----	-AAC	-----	-----GT	GGAGAAGGTG	CGCTTCATCC	CGATCTCCGG	[584]
083_EF1a_Trypanosoma_rangeli	-----	-----	-----	-----	-----	-AAC	-----	-----GT	GGAGAAGGTG	CGCTTCGTCC	CCATCTCCGG	[584]
084_EF1a_Ploeotia_costatum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCCTTCATCC	CCATCTCTGG	[581]
085_EF1a_Ploeotia_oblonga_cDNA	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCCTTCATCC	CCATCTCTGG	[581]
086_EF1a_Ploeotia_oblonga_gDNA	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCCTTCGTGC	CGATCTCCGG	[581]
087_EF1a_Entosiphon_sp.	-----	-----	-----	-----	-----	-AAT	-----	-----CC	CGACAAAGTT	AACCTTCATCC	CCATCTCCGG	[581]
088_EF1a_Entosiphon_sulcatum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGACAAAGTT	AACCTTCATCC	CGATCTCTGG	[509]
089_EF1a_Peranema_tricophorum	-----	-----	-----	-----	-----	-AAT	-----	-----CC	AGAAAAAGTA	CCCTTCATCC	CAATCAGTGG	[581]
090_EF1a_P. tricophorum_FJ807243	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAAAAAGTT	CCCTTCATCC	CAATCAGTGG	[575]
091_EF1a_Colacium_elongatum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGATAAGGTT	CCCTTCATCC	CGATCTCTGG	[581]
092_EF1a_Colacium_mucronatum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCGTTTCATCC	CGATCTCTGG	[527]
093_EF1a_Colacium_sp.	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCCTTCATCC	CGATCTCTGG	[581]
094_EF1a_Colacium_vesiculosum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCGTTTCATCC	CGATCTCTGG	[539]
095_EF1a_Cryptoglena_pigra	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGATAAGGTT	CCGTTTCATCC	CGATCTCTGG	[581]
096_EF1a_Cryptoglena_skujai	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGATAAGGTT	CCATTCATCC	CGATCTCTGG	[581]
097_EF1a_Discoplastis_adunca	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCCTTCATCC	CGATCTCTGG	[581]
098_EF1a_Euglena_clara	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCCTTCATCC	CGATCTCCGG	[461]
099_EF1a_Euglena_deses	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGAAAAAGTT	CCATTCATCC	CGATCTCTGG	[581]
100_EF1a_Euglena_exilis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGAAAAAGTT	CCCTTCATCC	CGATCTCTGG	[461]
101_EF1a_Euglena_geniculata	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCCTTCATCC	CGATCTCTGG	[581]
102_EF1a_Euglena_gracilis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCCTTCATCC	CGATCTCTGG	[581]
103_EF1a_E. gracilis_X16890	-----	-----	-----	-----	-----	-AAT	-----	-----CC	GGAAAAAGTT	CCCTTCATCC	CGATCTCTGG	[584]
104_EF1a_E. gracilis_FJ807238	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCCTTCATCC	CGATCTCTGG	[584]
105_EF1a_Euglena_granulata	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGACAAAGTT	CCCTTCATCC	CGATCTCTGG	[581]
106_EF1a_Euglena_hiemalis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCCTTCATCC	CGATCTCTGG	[581]
107_EF1a_Euglena_longa	-----	-----	-----	-----	-----	-AAC	-----	-----CT	TGAGAAGGTT	CGCTTCATCC	CGATCTCTGG	[98]
108_EF1a_Euglena_stellata	-----	-----	-----	-----	-----	-AAT	-----	-----CC	TGAGAAGGTT	CCCTTCATCC	CGATCTCTGG	[461]
109_EF1a_Euglena_viridis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAAAAAGTT	CCCTTCATCC	CGATCTCTGG	[539]
110_EF1a_Eutreptia_pertyi	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGACAAAGTT	CCCTTCATCC	CGATCTCTGG	[581]
111_EF1a_Eutreptiella_pomquetensis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCCTTCATCC	CGATCTCTGG	[191]
112_EF1a_Lepocinclis_acus_var._major	-----	-----	-----	-----	-----	-AAT	-----	-----CC	TGACAAAGTT	CCATTTATCC	CGATCTCTGG	[581]
113_EF1a_Lepocinclis_ovum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	AGACAAAGTT	CCATTCATCC	CGATCTCTGG	[461]
114_EF1a_Lepocinclis_spyrogyroides_clone_38	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
115_EF1a_L. spyrogyroides_clones_1_&_4	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCATTCATCC	CGATCTCTGG	[581]
116_EF1a_Lepocinclis_steinii	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAAAAAGTT	CCATTCATCC	CGATCTCTGG	[461]
117_EF1a_Lepocinclis_triapteris	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGAAAAAGTT	CCCTTCATCC	CGATCTCCGG	[581]
118_EF1a_Monomorphina_aenigmatica	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCGTTTCATCC	CGATCTCTGG	[581]
119_EF1a_Monomorphina_pyrum	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCATTCATCC	CGATCTCTGG	[581]
120_EF1a_Monomorphina_sp.	-----	-----	-----	-----	-----	-AAT	-----	-----CC	TGAGAAGGTT	CCCTTCATCC	CGATCTCTGG	[581]
121_EF1a_Phacus_orbicularis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGACAAAGTT	CCCTTCATCC	CGATCTCTGG	[539]
122_EF1a_Phacus_similis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGAGAAGGTT	CCGTTTCATCC	CGATCTCCGG	[581]
123_EF1a_Strombomonas_acuminata	-----	-----	-----	-----	-----	-AAT	-----	-----CC	TGACAAAGTT	CCCTTCATCC	CGATCTCTGG	[584]
124_EF1a_Strombomonas_borysteniensis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCGTTTCATCC	CGATCTCCGG	[461]
125_EF1a_Strombomonas_napiformis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	CGACAAAGTT	CCATTCATCC	CGATCTCTGG	[581]
126_EF1a_Strombomonas_ovalis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGACAAAGTT	CCATTCATCC	CGATCTCTGG	[581]
127_EF1a_Strombomonas_triquetra	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGACAAAGTT	CCATTCATCC	CGATCTCTGG	[581]
128_EF1a_Trachelomonas_conradi	-----	-----	-----	-----	-----	-AAC	-----	-----CC	GGAGAAGGTT	CCGTTTCATCC	CGATCTCTGG	[581]
129_EF1a_Trachelomonas_echinata	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGAGAAGGTT	CCCTTCATCC	CGATCTCTGG	[581]
130_EF1a_Trachelomonas_lefevrei	-----	-----	-----	-----	-----	-TAC	-----	-----CC	TGACAAAGTT	CCATTCATCC	CGATCTCCGG	[581]
131_EF1a_Trachelomonas_planctonica	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGACAAAGTT	CCATTCATCC	CGATCTCTGG	[461]
132_EF1a_Trachelomonas_similis	-----	-----	-----	-----	-----	-AAC	-----	-----CC	TGAGAAGGTT	CCATTCATCC	CGATCTCTGG	[461]

	670	680	690	700	710	720	730	740	750	760	770	
133_EFL1a_Trachelomonas_volvocina	-----	-----	-----	-----	-----	-----AAC-----	-----	-----CC	CGACAAGGTG	CCTTTTCATCC	CCATTCTCTGG	[581]
134_EFL_environmental_sample_EU810232	-----	-----	-----	-----	-----	-----	-----	-----GC	CAAGAGCGTG	CCCGTCATCC	CCATCTCCGG	[617]
135_EFL_Creolimax_fragrantissima	-----	-----	-----	-----	-----	-----	-----	-----GC	CAAGTCTGTG	CCCATCATCC	CCATCTCTGG	[545]
136_EFL_Sphaeroforma_arctica	-----	-----	-----	-----	-----	-----	-----	-----GC	CAAGTCCGTA	CCCATCATGC	CCATTTCAGG	[617]
137_EFL_Monosiga_brevicollis	-----	-----	-----	-----	-----	-----	-----	-----GA	GGAGTCCGTC	CCCGTCTCC	CCATCTCGGG	[617]
138_EFL_Xenopus_laevis	-----	-----	-----	-----	-----	-----	-----	-----CT	CAAGTCCGTG	CCCATCATCC	CCATCTCCGG	[617]
139_EFL_Spizellomyces_punctatus	-----	-----	-----	-----	-----	-----	-----	-----CA	GGGACGCGTT	CCCGTCATCC	CCATCTCTGG	[620]
140_EFL_Gaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----CA	GGGCTGTGTC	CCCGTCATCC	CCATCTCTGG	[323]
141_EFL_Allomyces_macrognus	-----	-----	-----	-----	-----	-----	-----	-----GA	GGGTTCCGTC	CCGGTCATCC	CGATCTCGGG	[620]
142_EFL_Blastocladiella_emersonii	-----	-----	-----	-----	-----	-----	-----	-----GA	GGAGTCCGTC	CCCGTCATCC	CTATCTCGGG	[620]
143_EFL_Catenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----GA	GCAATCCGTC	CCCGTCATCC	CCATCTCCGG	[323]
144_EFL_Coelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----GA	AAATGCGTT	CCCATCTCC	CCATCTCTGG	[323]
145_EFL_Entomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----GA	GCAGAACGTT	CCCGTCTCC	CCATCTCCGG	[320]
146_EFL_Basidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----GA	ACAATCCGTT	CCCGTTATYC	CTATCTCCGG	[320]
147_EFL_Conidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----AA	GGACTCCGTC	CCCATTCTCC	CCATCTCTGG	[320]
148_EFL_Tetraselmis_tetratele	-----	-----	-----	-----	-----	-----	-----	-----CT	CGGCTGCGTG	CCCATCTGCG	CCATCTCTGG	[566]
149_EFL_Mesostigma_viride	-----	-----	-----	-----	-----	-----	-----	-----GC	GGAGAGCGTC	CCCATCATCC	CGATTTCGGG	[620]
150_EFL_Ostreococcus_tauri	-----	-----	-----	-----	-----	-----	-----	-----GA	AAAGTCCGTC	CCGGTCTCC	CGATCTCCGG	[617]
151_EFL_Nephroselmis_olivacea	-----	-----	-----	-----	-----	-----	-----	-----AT	GGACTCAGTG	CCCGTCATGC	CCATCTCCGG	[593]
152_EFL_Micromonas_pusilla	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGTGCCTG	CCCGTCATCC	CAATCTCCGG	[566]
153_EFL_Parachlorella_kessleri	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGTCTGTG	CCCATCATCC	CCATCTCCGG	[575]
154_EFL_Helicospiridium_sp.	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGTGCCTG	CCCGTCATCC	CTATCTCTGG	[617]
155_EFL_Gonium_pectorale	-----	-----	-----	-----	-----	-----	-----	-----AA	GGACTCCGTC	CCCGTCTGCG	CCATCTCGGG	[617]
156_EFL_Chlamydomonas_reinhardtii	-----	-----	-----	-----	-----	-----	-----	-----AA	CAAGTCCGTC	CCCATCTGCG	CCATCTCCGG	[617]
157_EFL_Scenedesmus_obliquus	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGAGCGTG	CCCGTCTGCG	CCATCTCCGG	[617]
158_EFL_“Oryza_sativa”	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGAGCGTG	CCCATCATCC	CCATCTCCGG	[617]
159_EFL_Chlorococcum_sp.	-----	-----	-----	-----	-----	-----	-----	-----GA	GGCCAGCGTG	CCCTTCATGC	CCATCTCTGG	[566]
160_EFL_Ochlochaete_hystrix	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGTGTGTG	CCCATCTGCG	CCATCTCTGG	[575]
161_EFL_Ulva_intestinalis	-----	-----	-----	-----	-----	-----	-----	-----GA	GGGAAGCGTG	CCATTCTGCG	CCATCTCTGG	[566]
162_EFL_Urospora_sp.	-----	-----	-----	-----	-----	-----	-----	-----GA	CAACAGCGTG	CCCTTCCTGCG	CCATCTCCGG	[419]
163_EFL_Gymnochlorella_stellata_ppc	-----	-----	-----	-----	-----	-----	-----	-----GA	TAAGTCTATA	CCCTTCTTAC	CTCATCTGCG	[653]
164_EFL_Lotharella_amoeboformis_ppc	-----	-----	-----	-----	-----	-----	-----	-----GA	AAAGTCTGTT	CCGTTCTGCG	CCCTCTCTGG	[398]
165_EFL_Chlorarachnion_reptans_ppc	-----	-----	-----	-----	-----	-----	-----	-----GA	GAATCTGTA	CCTTTCTCTCC	CTCTCAGTGG	[425]
166_EFL_Chondrus_crispus	-----	-----	-----	-----	-----	-----	-----	-----GC	CAAGAGCGTG	CCSGTCATCC	CGATMTCTGG	[617]
167_EFL_Gracilaria_changii	-----	-----	-----	-----	-----	-----	-----	-----GC	CAAGAGTGTG	CCGGTCTGTT	CGATTTCGGG	[617]
168_EFL_Cryptomonas_ovata	-----	-----	-----	-----	-----	-----	-----	-----AT	GGACTCTGTT	CCGGTCTGTT	CGATCTCGGG	[572]
169_EFL_Guillardia_theta	-----	-----	-----	-----	-----	-----	-----	-----GC	TGCTTCCGTC	CCCGTCATCC	CCATCTCCGG	[614]
170_EFL_Rhodomonas_salina	-----	-----	-----	-----	-----	-----	-----	-----GC	CAAGTCCGTC	CCCATCGTCC	CTATCTCCGG	[614]
171_EFL_Pavlova_lutheri	-----	-----	-----	-----	-----	-----	-----	-----GA	GTCGGGCGTG	CCCGTCTGCG	CCATCTCGGG	[617]
172_EFL_Chrysochromulina_sp.	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGTCGACC	CCGGTCTCC	CGCTCTCGGG	[437]
173_EFL_Isochrysis_galbana	-----	-----	-----	-----	-----	-----	-----	-----GA	GCAATCAACG	CCGGTGCTCC	CGATCTCGGG	[617]
174_EFL_Emiliania_huxleyi	-----	-----	-----	-----	-----	-----	-----	-----GC	CCAGTCCGTC	CCCGTCTGCG	CGCTCTCCGG	[617]
175_EFL_Achnanthes_kuwaitensis	-----	-----	-----	-----	-----	-----	-----	-----CC	GAAGAAGGTC	CCCTTTATCC	CTTACTCCGG	[536]
176_EFL_Asterionella_glacialis	-----	-----	-----	-----	-----	-----	-----	-----CC	CAAGAAGGTT	CCCTTCATTC	CTTACTCTGG	[536]
177_EFL_Detionula_confervacea	-----	-----	-----	-----	-----	-----	-----	-----CC	CAAGAAGGTC	CCTTTTCATCC	CTTACTCTGG	[536]
178_EFL_Thalassiosira_pseudonana	-----	-----	-----	-----	-----	-----	-----	-----CC	CAAGAAGGTT	CCTTTTCATCC	CCTATTCCGG	[599]
179_EFL_Perkinsus_marinus	-----	-----	-----	-----	-----	-----	-----	-----GA	GAATGTGTT	CCCGTCATTC	CTATTCTCTGG	[626]
180_EFL_Oxyrrhis_marina	-----	-----	-----	-----	-----	-----	-----	-----GA	CACGAATGTT	CCGTTTCATGC	CCATCTCTGG	[623]
181_EFL_Karlodinium_micrum	-----	-----	-----	-----	-----	-----	-----	-----GA	AAAGAATACT	CCTTACATGC	CAATCTCTGG	[623]
182_EFL_Amphidinium_carterae	-----	-----	-----	-----	-----	-----	-----	-----AA	GAAGAACACA	CCGATGCTTC	CAATCTCAGG	[623]
183_EFL_Heterocapsa_triquetra	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGAACACG	CCGGTCTGCG	CGATCTCTGG	[623]
184_EFL_Planoglabratella_operculis	-----	-----	-----	-----	-----	-----	-----	-----AC	AAAGAAGATC	CCATTTATTC	CATTGTCTGG	[443]
185_EFL_Reticulomyxa_filosa	-----	-----	-----	-----	-----	-----	-----	-----AC	CAAAAAAATC	CCATTTATTC	CAATGAGTGG	[533]
186_EFL_Lotharella_amoeboformis	GAGAAG----	-----	-----	-----	-----GGAAAGAA	GAAGGGC----	-----	-----CC	ACCTATGATC	CCGGTATATCC	CTATCTCTGG	[686]
187_EFL_Gymnochlorella_stellata	AAGAAGGACG	ACAAGGGTGG	CAAGAAGAAA	GGAAAGGGAG	AGAAGAAAAA	GGCCGGA----	-----	-----CC	TAACCTCATT	CCCGTAATTC	CTATCTCTGG	[638]
188_EFL_Chlorarachnion_reptans	ACGGAC----	-----	-----	-----	-----AAAAAGAA	GAAGGGA----	-----	-----CC	AAACCTCGTT	CTGTGTTATCC	CTATTCTCTGG	[569]
189_EFL_Diplonema_ambulator	-----	-----	-----	-----	-----	-----	-----	-----GA	GAACAGCGTC	CCCGTCTCC	CCATCTCCGG	[218]
190_EFL_Diplonema_papillatum	-----	-----	-----	-----	-----	-----	-----	-----GA	GCAGTGCGTG	CCGATCGTCC	CGATCTCTGG	[590]
191_EFL_Rhynchopus_euleeides	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGTCCGTC	CCCGTCTGCG	CCATCTCGGG	[230]
192_EFL_Petalomonas_cantuscygni	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGTCCGTC	CCCGTCATCC	CCATCTCCGG	[461]
193_EFL_Trypanoplasma_borrel	-----	-----	-----	-----	-----	-----	-----	-----GA	GAACGAGATC	CCCATCATCC	CCATGGCTGG	[635]
194_EFL_Neobodo_salien	-----	-----	-----	-----	-----	-----	-----	-----GA	GAAGGAGATC	CCGATCTCC	CGCTCGCCGG	[581]

	780	790	800	810	820	830	840	850	860	870	880	
001_aEF1a_Desulfurococcus_mobilis	TTGGACCGGT	GAGAACCTGA	TAGAGCGCTG	GCCT-----	-----AACA	TGCCATGG--	----TACAAT	GGT-----	-----	-----	-----	[642]
002_aEF1a_Sulfolobus_solfataricus	ACCTGCAAGG	GATAACATGA	CTCATAGATC	AGAA-----	-----ATA	TGAATATG--	----TATAAC	GGT-----	-----	-----	-----	[636]
003_aEF1a_Thermoplasma_acidophilum	cTaCaAgGga	gACaAcgTaA	cAAAgCcCaG	cCCG-----	-----aTa	tGcCaTgG--	----tACaAg	gGc-----	-----	-----	-----	[618]
004_aEF1a_Halobacterium_sp.	GTTCGAGGGC	GACAACGTCT	CGGACCACCT	GGAC-----	-----AACA	CACCGTGG--	----TACGAC	GGC-----	-----	-----	-----	[609]
005_aEF1a_Methanocaldococcus_jannaschii	ATTGAAGGGA	GACAACGTGC	TTAAAAGATC	AGAA-----	-----AACA	TGCCATGG--	----TACAAA	GGT-----	-----	-----	-----	[624]
006_EF1a_Acanthamoeba_culbertsoni	TTTCACGGGT	GACAACATGG	TGGACAAGAC	CGAC-----	-----AAGA	TGCCCTGG--	----TACAAG	GGA-----	-----	-----	-----	[606]
007_EF1a_Neoparamoeba_branchipila	ATGGAACGGT	GACAACATGT	TGGAGCGCTC	CACC-----	-----AACA	TGAACCTGG-	----TACAAG	GGT-----	-----	-----	-----	[591]
008_EF1a_Planoprotostelium_aurantium	ATGGAACGGa	GACAACATGC	TCGAGAAGCC	CGAC-----	-----AACA	TGCCCTGG--	----TGAAG	GGA-----	-----	-----	-----	[588]
009_EF1a_Physarum_polycephalum	ATGGAACGGa	GACAACATGC	TCGAGAAGTC	CGCC-----	-----AACC	TCCCATGG--	----TACAAG	GGA-----	-----	-----	-----	[588]
010_EF1a_Amoebidium_parasiticum	CTTCAACGGT	GACAACATGG	TCGAGCCAC	CGAC-----	-----AACA	TGCCCTGG--	----TACAAG	GGATGGGAGG	TGGAGCGCAa	GGAGGGC--	-----	[627]
011_EF1a_Capsaspora_owczaraki	CTGGCACGGC	GACAACATGC	TCGAGGCCCT	CGAG-----	-----AACA	TGCCCTGG--	----TTCAAG	GGATGGACTA	TCGAGCGCAa	GGAGGGC--	-----	[657]
012_EF1a_Monosiga_ovata	CTGGCACGGC	GACAACATGA	TTGAGGCCCT	GGAG-----	-----AAGC	TCCCTTGG--	----TACAAG	GGCTGGGAGA	TTACCCGCAa	GGATGGC--	-----	[636]
013_EF1a_Salpingoeca_amphoridium	CTGGCACGGC	GACAACATGA	TTGAGGCCCT	GGAG-----	-----AAGC	TGCCCTGG--	----TACAAG	GGCTGGGAGG	TCCAGCGCAa	GGATGAT--	-----G	[631]
014_EF1a_Ceanorhabditis_elegans	ATTCAACGGa	GACAACATGC	TTGAGGTCTC	CTCC-----	-----AACA	TGCCATGG--	----TTCAAG	GGATGGGCCG	TTGAGCGCAa	GGAGGGA--	-----	[663]
015_EF1a_Drosophila_melanogaster	ATGGCACGGC	GACAACATGT	TGGAACCCCT	TACC-----	-----AACA	TGCCCTGG--	----TTCAAG	GGATGGGAAGG	TGGAGCGCAa	GGAGGGT--	-----	[663]
016_EF1a_Ornithorhynchus_anatinus	TTGGCACGGa	GACAACATGT	TGGAACCCCT	TCCT-----	-----AACA	TGCTTGG--	----TTTAAG	GGCTGGGAAGG	TAGAGCGGAa	GGAAAGC--	-----	[663]
017_EF1a_Xenopus_laevis	ATGGCATGGA	GACAACATGC	TGGAAGGTAG	CACC-----	-----ATA	TGCCCTGG--	----TTTAAa	GGCTGGGAAGA	TTGAAGGGAa	AGAAGGA--	-----	[663]
018_EF1a_Saccharomyces_cerevisiae	TTGGAACGGT	GACAACATGA	TTGAAGCTAC	CACC-----	-----AACG	CTCCATGG--	----TACAAG	GGTTGGGAAA	AGGAACCAa	GGCCGGT--	-----	[657]
019_EF1a_Neuropora_crassa	CTTCAACGGC	GACAACATGC	TTGAGCCCTC	CACC-----	-----AACC	TGCCCTGG--	----TACAAG	GGTTGGGAGA	AGGAGACCAa	GGCCGGC--	-----	[660]
020_EF1a_Cryptococcus_neoformans	TTGGCACGGT	GACAACATGT	TGGAGGAGAC	CACC-----	-----AACA	TGCCCTGG--	----TACAAG	GGATGGACCA	AGGAGACCAa	GTCTGGT--	-----	[657]
021_EF1a_Schizophyllum_comune	CTGGCACGGC	GACAACATGT	TGGAGGAGTC	CACC-----	-----AACA	TGCCGTGG--	----TACAAG	GGCTGGACCA	AGGAGACCAa	GGCTGGT--	-----	[657]
022_EF1a_Monoblepharis_insignis	TTGGAACGGT	GACAACATGA	TTGAGGCCAC	CACC-----	-----AACA	TGCCCTGG--	----TACAAG	GGCTGGGAGA	AGGAGACCAa	GTCCGGC--	-----	[543]
023_EF1a_Rhizophyctis_harderi	ATGGCACGGT	GACAATATGT	TGGAGCCTTC	TGAA-----	-----AACA	CCCCATGG--	----TTCAAG	GGATGGACCA	AGGAACCAa	GGCTGGA--	-----	[543]
024_EF1a_Harpochytrium_sp.	ATGGCACGGa	GACAACATGA	TTGAGGCCCT	GGAC-----	-----AACA	TGCCGTGG--	----TTCAAG	TCATGGCAGA	AGGAGACCAa	AGGTGGA--	-----	[495]
025_EF1a_Absidia_glauca	CTGGCACGGT	GACAACATGT	TGGAAGAATC	TACC-----	-----AACA	TGCCCTGG--	----TACAAG	GGATGGAACA	AGGAACCAa	GGCTGGT--	-----	[657]
026_EF1a_Smittium_simulii	TTTCATGGT	GATAACATGA	TTGAAGCTTC	TACT-----	-----AACA	TGCTTGG--	----TATAAG	GGATGGACTA	AAGAAACCAa	ATCTGGT--	-----	[630]
027_EF1a_Chara_australis	GTTCGAGGGG	GATAATATGA	TCGAGAGGTC	GACG-----	-----ATA	TGGGGTGG--	----TACAAA	GGC-----	-----	-----	-----	[588]
028_EF1a_Physcomitrella_patens	GTTTCAGGGC	GACAACATGA	TCGAGAGGTC	GACG-----	-----AACC	TGGAGTGG--	----TACGAG	GGC-----	-----	-----	-----	[639]
029_EF1a_Acetabularia_acetabulum	TTTCAGGGa	GACAACATGC	TTGAGAAGAG	TAGC-----	-----AACA	TGAACCTGG-	----TACAAG	GGA-----	-----	-----	-----	[588]
030_EF1a_Chaetomorpha_coliformis	CTGGAATGGT	GATAATATGT	TGAGCGTTTC	GTCC-----	-----AACA	TGGGCTGG--	----TACAAG	GGC-----	-----	-----	-----	[438]
031_EF1a_Arabidopsis_thaliana	ATTTCAGGGa	GACAACATGA	TTGAGAGGTC	AACC-----	-----AACC	TTGACTGG--	----TACAAG	GGA-----	-----	-----	-----	[639]
032_EF1a_Oryza_sativa	GTTTCAGGGT	GACAACATGA	TTGAGAGGTC	CACC-----	-----AACC	TTGACTGG--	----TACAAG	GGC-----	-----	-----	-----	[639]
033_EF1a_Cyanidioschyzon_merolae	CTGGACTGGC	GACAACCTGT	TCGAGCGAGT	GCCATCGGAC	CATCCGCTTG	CCAAGTGG--	----TACAAG	GGa-----	-----	-----	-----	[651]
034_EF1a_Porphyrha_jezoensis	CTGGACGGGT	GAGAACCTGT	TCGAGAGGAC	CGATAAGACG	CACGCCCTCG	GCAAGTGG--	----TACAAG	GGC-----	-----	-----	-----	[651]
035_EF1a_Plasmodium_falciparum	TTTTGAAGGT	GATAATTTAA	TGGAAAATC	AGAT-----	-----AAAA	CACCATGG--	----TATAAA	GGA-----	-----	-----	-----	[633]
036_EF1a_Cryptosporidium_parvum	TTTCGTTGGT	GATAATATGG	TTGAGAGATC	TGAC-----	-----AAGA	TGCCATGG--	----TATAAG	GGA-----	-----	-----	-----	[627]
037_EF1a_Tetrahymena_pyriformis	TTTCAATGGT	GACAACATGT	TGGAAGATC	CACC-----	-----AACG	CCCCCTGG--	----TACAAG	GGT-----	-----	-----	-----	[642]
038_EF1a_Stentor_coeruleus	TTGGCTCGGT	GATAACATGG	TTGAACGCTC	ACCA-----	-----AACC	TTCTTGG--	----TACAAT	GGC-----	-----	-----	-----	[594]
039_EF1a_Phytophthora_parasitica	CTGGGAGGGT	GACAACATGC	TGACGCCCTC	CACC-----	-----AACA	TGCCGTGG--	----TACAAG	GGC-----	-----	-----	-----	[639]
040_EF1a_Mallomonas_rasilis	CTGGGTTGGT	GACAATATGA	TCGATCGCTC	TMCA-----	-----ATA	TGGCTTGG--	----TACAAG	GGT-----	-----	-----	-----	[600]
041_EF1a_Phaeodactylum_tricornutum	TTGGGAGGGT	GACAACATGG	TGGAAAAGTC	CACT-----	-----AACA	TGGCTTGG--	----TACAAG	GGa-----	-----	-----	-----	[639]
042_EF1a_Thalassiosira_pseudonana	CTGGTCTGGC	GAGAACCTCA	CTGAACGAAC	CAAA-----	-----AACA	TGGCTTGG--	----TACAAT	GGT-----	-----	-----	-----	[639]
043_EF1a_Malawimonas_jakobiformis	CTGGAACGGT	GACAACATGA	TGGAGAAGTC	TCCC-----	-----AACA	TGCCGTGG--	----TACAAG	GGC-----	-----	-----	-----	[570]
044_EF1a_Malawimonas_californiana	CTGGCTCGGC	GACAACATGA	TCGAGCGCTC	TCCC-----	-----AACA	TGGGCTGG--	----TACAAG	GGC-----	-----	-----	-----	[639]
045_EF1a_Trimastix_pyriformis	CTTCACGGC	GACAACATGC	TGGAACCCCT	CGCC-----	-----AACA	TGCCCTGG--	----TGAAG	GGC-----	-----	-----	-----	[597]
046_EF1a_Dinenympha_exilis	CTTCAATGGT	GACAACATGC	TCGATCGATC	TACG-----	-----AACA	TGCCGTGG--	----TATAAG	GAC-----	-----	-----	-----	[579]
047_EF1a_Monocercomonoides_sp.	ATTCCAAGGA	GATAACATGC	TTGAAAGATC	AGCA-----	-----AACA	TGCCATGG--	----TACAAG	GGA-----	-----	-----	-----	[639]
048_EF1a_Streblomastix_strix	GTTCTGCGGA	GATAACATGC	TTGAACACTC	ACCA-----	-----AACA	TGCCATGG--	----TACAAG	GGA-----	-----	-----	-----	[588]
049_EF1a_Histomonas_meleagridis	ATTCTGCGGT	GATAACATGA	CCGAAAAGTC	ACCA-----	-----AACA	TGCCATGG--	----TACAA	GGC-----	-----	-----	-----	[639]
050_EF1a_Trichomonas_vaginalis	CTGGGCTGGC	GACAACATGA	CAGAGAAGTC	TCCA-----	-----AACA	TGCCATGG--	----TACAAT	GGC-----	-----	-----	-----	[639]
051_EF1a_Hexamita_inflata	TTGGACTGGT	GATAACATGA	TGGAAGCCTC	TCCA-----	-----AAGA	CTCCATGG--	----TACAAG	GGC-----	-----	-----	-----	[591]
052_EF1a_Spironucleus_vortens	CTGGACTGGT	CTCAACATCA	TGGAGAAGGA	CACC-----	-----GAGA	AGACCTGG--	----TATGAC	GGC-----	-----	-----	-----	[591]
053_EF1a_Andalucia_incarcerata	ATGGGTGCGT	GACAACATGC	TTGAGCGCTC	TCCC-----	-----AACA	TGCCCTGG--	----TGAAG	GGT-----	-----	-----	-----	[588]
054_EF1a_Seculamonas_ecuadoriensis	CTGGCACGGC	GACAACATGC	TTGAGCGCTC	GAAg-----	-----AACA	GTGCTGG--	----TACACC	GGC-----	-----	-----	-----	[639]
055_EF1a_Reclinomonas_americana	CTGGAACGGC	GACAACATGC	TTGAGCGCTC	GCCC-----	-----AACA	GTGCTGG--	----TACAAG	GGT-----	-----	-----	-----	[639]
056_EF1a_Histiona_aroides	CTGGAACGGC	GACAACATGT	TGGAGAAGTC	TGCC-----	-----AACA	TGCCATGG--	----TACAAG	GGC-----	-----	-----	-----	[639]
057_EF1a_Jakoba_libera	CTGGAACGGC	GACAACATGT	TGGAGCGCTC	CAAG-----	-----AACA	CCCCCTGG--	----TACACT	GGA-----	-----	-----	-----	[639]
058_EF1a_Stachyamoeba_lipophora	TTTCAACGGT	GATAACATGA	TGGATCGTTT	TACC-----	-----AACA	TGCCCTGG--	----TACAAG	GGA-----	-----	-----	-----	[636]
059_EF1a_Naegleria_gruberi	TTGGACTGGT	GACAACATGA	TTGAAAAGTC	TGAC-----	-----AAGA	TGGGTTGG--	----TATAAG	GGT-----	-----	-----	-----	[588]
060_EF1a_Naegleria_andersoni	TTGGACTGGT	GATAACATGA	TTGAAAAGTC	TGAC-----	-----AAGA	TGCCATGG--	----TACAAG	GST-----	-----	-----	-----	[606]
061_EF1a_Acrasis_rosea	ATGGGCGGGT	GACAACATGT	TGGAGAATAT	CACC-----	-----AACA	TGCCATGG--	----TACAAG	GGA-----	-----	-----	-----	[588]
062_EF1a_Sawyeria_marylandensis	ATGGACAGGA	GAAAACATGA	TTGAAAATAT	AGAT-----	-----AACT	TAAAATGG--	----TATACT	GGT-----	-----	-----	-----	[639]
063_EF1a_Bodo_saltans	TTGGGAAGGC	GACAACATGA	TCGAGAAGTC	CTCC-----	-----CGCA	TGGAAGTGG--	----TACAAG	GGC-----	-----	-----	-----	[579]
064_EF1a_Rhynchobodo_sp.	CTGGAACGGY	GACAACATGC	TTGAGTCTCT	CACC-----	-----AACA	TGCCCTGG--	----TACAAG	GGA-----	-----	-----	-----	[573]
065_EF1a_Rhynchomonas_nasuta	CTGGAACGGC	GACAACATGA	TCGAGAAGTC	CGAC-----	-----AACA	TGCCGTGG--	----TACAAG	GGC-----	-----	-----	-----	[636]
066_EF1a_Dimastigella_trypaniformis	CTGGAACGGT	GACAACATGC	TGGAGAAGTC	CGCG-----	-----AACA	TGCCGTGG--	----TACAAG	GGC-----	-----	-----	-----	[594]

	780	790	800	810	820	830	840	850	860	870	880	
067_EF1a_Perkinsiella_amoebae	CTGGATGGGC	GACAACATGA	TTGAGCGCTC	CACC-----	-----AACA	TGAAGTGG--	----TACAAG	GGC-----	-----	-----	-----	[564]
068_EF1a_Leishmania_major	CTGGCAGGGC	GACAACATGA	TCGAGAAGTC	GGAC-----	-----AACA	TGCCGTGG--	----TACAAG	GGT-----	-----	-----	-----	[639]
069_EF1a_Leishmania_infantum	CTGGCAGGGC	GACAACATGA	TCGAGAGGTC	GGAC-----	-----AACA	TGCCGTGG--	----TACAAG	GGT-----	-----	-----	-----	[639]
070_EF1a_Leishmania_braziliensis	CTGGCAGGGC	GACAACATGA	TCGACAAGTC	GGAG-----	-----AGCA	TGGCGTGG--	----TACAAG	GGT-----	-----	-----	-----	[639]
071_EF1a_Leishmania_tarentolae	CTGGCAGGGC	GACAACATGA	TCGAGAAGTC	GGAC-----	-----AACA	TGCCGTGG--	----TACAAG	GGT-----	-----	-----	-----	[639]
072_EF1a "L. tarentolae" FJ807257	GTGGAACGGT	GACAATATGA	TCGAGAAGTC	GGAG-----	-----AAGA	TGCCGTGG--	----TACAAG	GGC-----	-----	-----	-----	[564]
073_EF1a_Leptomonas_costaricensis	CTGGCAGGGC	GACAACATGA	TCGACAAGTC	GGAC-----	-----AACA	TGTCGTGG--	----TACAAG	GGG-----	-----	-----	-----	[387]
074_EF1a_Leptomonas_bifurcata	CTGGCAGGGC	GACAACATGA	TCGAGAAGTC	CGAC-----	-----AACA	TGGCGTGG--	----TACAAG	GGT-----	-----	-----	-----	[399]
075_EF1a_Leptomonas_podlipaevi	CTGGCAGGGT	GACAACATGA	TCGAGAAGTC	CGAC-----	-----AACA	TGCCGTGG--	----TACAAG	GGT-----	-----	-----	-----	[387]
076_EF1a_Herpetomonas_muscarum	CTTCCAGGGC	GACAACATGA	TCGACCGCTC	CACC-----	-----AACA	TGAGCTGG--	----TACAAG	GGC-----	-----	-----	-----	[552]
077_EF1a_Herpetomonas_pessoai	CTGGCAGGGC	GACAACATGA	TCGAGAAGTC	GACC-----	-----AACA	TGGGCTGG--	----TACAAG	GGT-----	-----	-----	-----	[420]
078_EF1a_Blastocrithidia_culicis	CTGGCTCGGC	GACAACATGA	TCGAGAAGTC	CGAC-----	-----AACA	TGTCCTGG--	----TACAAG	GGC-----	-----	-----	-----	[552]
079_EF1a_Trypanosoma_brucei	ATGGCAGGGC	GACAACATGA	TTGAGAAATC	CGAG-----	-----AAGA	TGCCATGG--	----TACAAG	GGT-----	-----	-----	-----	[639]
080_EF1a_T. brucei equiperdum	ATGGCAGGGC	GACAACATGA	TTGAGAAATC	CGAG-----	-----AAGA	TGCCATGG--	----TACAAG	GGT-----	-----	-----	-----	[561]
081_EF1a_Trypanosoma_aviium	ATGGCAGGGC	GACAACATGA	TCGAGAAGTC	CGAG-----	-----AAGA	TGCCGTGG--	----TACAAG	GGG-----	-----	-----	-----	[585]
082_EF1a_Trypanosoma_cruzi	CTGGCAGGGC	GACAACATGA	TCGACAAGTC	GGAA-----	-----AATA	TGCCGTGG--	----TACAAG	GGC-----	-----	-----	-----	[639]
083_EF1a_Trypanosoma_rangeli	GTGGCAGGGC	GACAACATGA	TCGAGAAGTC	CGAC-----	-----AATA	TGCCATGG--	----TACAAG	GGC-----	-----	-----	-----	[639]
084_EF1a_Ploetitia_costatum	CTGGAATGGT	GACAACATGA	TTGAGCCGTC	AGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGC-----	-----	-----	-----	[636]
085_EF1a_Ploetitia_oblonga_cDNA	CTGGAATGGT	GACAACATGA	TTGAGCCGTC	AGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGC-----	-----	-----	-----	[636]
086_EF1a_Ploetitia_oblonga_gDNA	CTGGGTGGT	GACAACATGT	TGGAGAAGTC	AGAC-----	-----AACA	TGGGATGG--	----TACAAG	GGC-----	-----	-----	-----	[636]
087_EF1a_Entosiphon_sp.	CTGGGTGGGA	GACAACATGA	TTGAGAAGTC	CGAC-----	-----AACA	TGGCCTGG--	----TACAAG	GGC-----	-----	-----	-----	[636]
088_EF1a_Entosiphon_sulcatum	CTGGGTGGGA	GACAACATGA	TTGAGAAGTC	CGAC-----	-----AACA	TGGCTTGG--	----TACAAG	GGT-----	-----	-----	-----	[564]
089_EF1a_Peranema_tricophorum	ATGGGTGGGA	GACAACATGA	TTGAAGCCAC	CGAA-----	-----AACA	TGCCATGG--	----TACAAA	GGG-----	-----	-----	-----	[636]
090_EF1a_P. tricophorum FJ807243	ATGGGTGGGA	GACAACATGA	TTGAAGCCAC	CGAR-----	-----AACA	TGCCATGG--	----TACAAR	GGG-----	-----	-----	-----	[630]
091_EF1a_Colacium_elongatum	ATGGAACGGT	GACAACATGT	TTGAGCCGTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
092_EF1a_Colacium_mucronatum	CTGGAATGGT	GACAACATGA	TTGATCCCTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGC-----	-----	-----	-----	[582]
093_EF1a_Colacium_sp.	ATGGAACGGT	GACAACATGA	TTGATCCCTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
094_EF1a_Colacium_vesiculosum	GTGGAATGGT	GACAACATGA	TTGATCCCTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGC-----	-----	-----	-----	[594]
095_EF1a_Cryptoglena_pigra	CTGGAATGGT	GACAACATGA	TTGAGCCCTC	GGAG-----	-----AACA	AGGGTTGG--	----TACAAG	GGC-----	-----	-----	-----	[636]
096_EF1a_Cryptoglena_skujai	CTGGAATGGT	GATAACATGA	TTGAGCCCTC	TGAG-----	-----AACA	TGGCTTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
097_EF1a_Discoplastis_adunca	GTGGAATGGT	GACAACATGA	TCGAGCCCTC	CGAC-----	-----AATA	TGACTTGG--	----TACAAG	GGC-----	-----	-----	-----	[636]
098_EF1a_Euglena_clara	TTGGAATGGT	GATAACATGA	TCGAAGCTTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[516]
099_EF1a_Euglena_deses	CTGGAACGGG	GACAACATGA	TTGAGCCCTC	TGAC-----	-----AACA	TGGGATGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
100_EF1a_Euglena_exilis	CTGGAACGGG	GACAACATGA	TTGAGGCGTC	CGAG-----	-----AATA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[516]
101_EF1a_Euglena_geniculata	CTGGAATGGT	GACAACATGA	TTGAGCCGTC	AGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
102_EF1a_Euglena_gracilis	CTGGAATGGT	GACAACATGA	TTGAGCCCTC	TGAT-----	-----AACA	TGGGCTGG--	----TACGAG	GGG-----	-----	-----	-----	[636]
103_EF1a E. gracilis X16890	CTGGAACGGG	GACAACATGA	TCGAGGCGTC	TGAA-----	-----AACA	TGGGATGG--	----TACAAG	GGG-----	-----	-----	-----	[639]
104_EF1a E. gracilis FJ807238	CTGGAATGGT	GACAACATGA	TTGAGCCCTC	TGAT-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[639]
105_EF1a_Euglena_stellata	CTGGAATGGG	GACAACATGA	TTGAGCCCTC	TGAT-----	-----AATA	TGTCTTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
106_EF1a_Euglena_hiemalis	CTGGAATGGT	GACAACATGA	TTGAGGCGTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
107_EF1a_Euglena_longa	CTGGAATGGC	GACAACATGA	TTGAGCCCTC	TGAC-----	-----AACA	TGGGATGG--	----TACAAG	GGG-----	-----	-----	-----	[153]
108_EF1a_Euglena_stellata	CTGGAATGGG	GACAACATGA	TTGAACCTTC	TGAG-----	-----AACA	TGGGTTGG--	----TACAAG	GGT-----	-----	-----	-----	[516]
109_EF1a_Euglena_viridis	ATGGAATGGT	GACAACATGA	TTGAGCCCTC	CGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGT-----	-----	-----	-----	[594]
110_EF1a_Eutreptia_pertyi	CTGGAATGGG	GACAACATGA	TTGAGCCCTC	TGAT-----	-----AATA	TGTCTTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
111_EF1a_Eutreptiella_pomquetensis	CTGGAATGGT	GACAACATGA	TTGAGCCGTC	AGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[246]
112_EF1a_Lepocinclis_acus_var._major	ATGGAATGGT	GACAATATGA	TTGAGGCGTC	CGAG-----	-----AACA	TGCCATGG--	----TATAAG	GGG-----	-----	-----	-----	[636]
113_EF1a_Lepocinclis_ovum	ATGGAACGGT	GACAATATGA	TTGATCCCTC	CGAC-----	-----AACA	TGCTTGG--	----TACAAG	GGG-----	-----	-----	-----	[516]
114_EF1a_Lepocinclis_spyrogyroides_clone_38												[0]
115_EF1a L. spyrogyroides clones_1_&_4	CTGGAATGGT	GATAACATGA	TCGACCCTTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
116_EF1a_Lepocinclis_steinii	CTGGAATGGG	GACAACATGA	TTGAGCCATC	AGAC-----	-----AACA	TGAAGTGG--	----TACAAG	GGC-----	-----	-----	-----	[516]
117_EF1a_Lepocinclis_triapteris	TTGGAATGGG	GACAACATGA	TTGAGCCCTC	TGAC-----	-----AACA	TGGGTTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
118_EF1a_Monomorphina_aenigmatica	CTGGAATGGG	GACAACATGA	TTGAGCCCTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
119_EF1a_Monomorphina_pyrum	CTGGAATGGT	GATAACATGA	TCGACCCTTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
120_EF1a_Monomorphina_sp.	GTGGAATGGT	GACAACATGA	TCGAGCCCTC	TGAA-----	-----AACA	TGGGTTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
121_EF1a_Phacus_orbicularis	CTGGAACGGT	GACAACATGA	TTGAGCCATT	TGAC-----	-----AACA	TGCCCTGG--	----TACAAG	GGG-----	-----	-----	-----	[594]
122_EF1a_Phacus_similis	CTGGAACGGT	GACAACATGA	TTGAGCCCTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGT-----	-----	-----	-----	[636]
123_EF1a_Strombomonas_acuminata	CTGGAACGGT	GACAACATGA	TTGAGGCTTC	TGAG-----	-----AATA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[639]
124_EF1a_Strombomonas_borysteniensis	ATGGAATGGT	GACAACATGA	TTGATCTCTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGT-----	-----	-----	-----	[516]
125_EF1a_Strombomonas_napiformis	TTGGAACGGG	GACAACATGA	TTGAGGCGTC	TGAG-----	-----AATA	TGACCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
126_EF1a_Strombomonas_ovalis	TTGGAATGGT	GACAACATGA	TTGAGGCGTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGT-----	-----	-----	-----	[636]
127_EF1a_Strombomonas_triquetra	CTGGAATGGT	GACAACATGA	TTGAAGCCTC	TGAG-----	-----AACA	TGCCCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
128_EF1a_Trachelomonas_conradi	GTGGAACGGT	GACAACATGA	TTGAGGCGTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
129_EF1a_Trachelomonas_echinata	TTGGAATGGG	GACAACATGA	TTGAGGCGTC	TGAG-----	-----AACA	TGGGATGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
130_EF1a_Trachelomonas_lefevrei	ATGGAATGGT	GACAATATGA	TCGAGCCCTC	TGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[636]
131_EF1a_Trachelomonas_plantonica	GTGGAATGGG	GACAACATGA	TCGAGCCCTC	CGAG-----	-----AACA	TGGGCTGG--	----TACAAG	GGG-----	-----	-----	-----	[516]
132_EF1a_Trachelomonas_similis	CTGGAATGGG	GACAACATGA	TTGAGGCTTC	TGAC-----	-----AACA	TGAATATGG--	----TACAAG	GGT-----	-----	-----	-----	[516]

	780	790	800	810	820	830	840	850	860	870	880	
133_EFL1a_Trachelomonas_volvocina	ATGGAATGGA	GACAACATGA	TTGAGGCGATC	TGAA-----	-----AACA	TGCGATGG--	-----TACAAG	GGA-----	-----	-----	-----	[636]
134_EFLenvironmental_sample_EU810232	ATACATGGGA	GACAACCTGA	TCAAGCCGAG	CACC-----	-----AACA	TGGACTGG--	-----TGGAAAG	GGAGTCGACG	TCAAGGTTGG	CGACAAG--	-----	[696]
135_EFLCreolimax_fragrantissima	ATGGTGTGGA	GACAACCTGA	TTACCAAGTC	CGAC-----	-----AAGA	TGGCTTGG--	-----TGGTCT	GGAGTTGATG	TCCTTGTGG	CAAGGAG--	-----	[624]
136_EFLSphaeroforma_arctica	ATGGTGTGGA	GATAACCTTA	TCACCAAGTC	CAAC-----	-----AACA	TGGATTGG--	-----TGGAAAG	GGAGCTGACG	TTCTTGTGCG	CAAGGAG--	-----	[696]
137_EFLMonosiga_brevicollis	CTGGAAGGGT	GATAACCTGA	TCACCAAGAGC	CACC-----	-----AACA	TGTCTTGG--	-----TGGAAAC	GGTGTGCGAC	TCAAGGTCGA	CAGCGAC--	-----	[696]
138_EFLXenopus_laevis	CTGGATCGGC	GACAACCTGC	TGACCAAGTC	CGCC-----	-----AAGA	TGCCGTGG--	-----TGGGAG	GGCGTGGATG	TCCTCGTCGG	CAAGGAG--	-----	[696]
139_EFLSpizellomyces_punctatus	ATGGATCGGT	GACAACCTGC	TCAAGAAGTC	TGAG-----	-----AAGA	TGTCTTGG--	-----TACCAA	GGTCAGGAGG	TTACCGCTCT	CTCTGGA--	-----	[699]
140_EFLGaertneriomyces_semiglobiferous	ATGGCAAGGA	GACAACCTGT	TGAAGAAGTC	CACC-----	-----AACA	TGCCCTGG--	-----TACTCT	GGTCAGGAGG	TCACCTGCCCT	CTCTGGC--	-----	[402]
141_EFLAllomyces_macrogynus	CTGGATGGGC	GACAACCTGC	TCAAGAAGTC	GGAG-----	-----AAGA	TGACGTGG--	-----TGGAAAT	GGCCAGAAGG	TCAAGTCGCT	CTCGGGC--	-----	[699]
142_EFLBlastocladiella_emersonii	CTGGATGGGC	GACAACCTGC	TCAAGCAGTC	GACC-----	-----AACA	TGACCTGG--	-----TGGGCT	GGCCAGGAGG	TCAAGCAGTC	CAACGGC--	-----	[699]
143_EFLCatenophylctis_sp.	CTGGATGGGC	GACAACCTGC	TCAAGCAATC	GACC-----	-----AACA	TGACCTGG--	-----TGGAAAG	GGTCAGAGG	TTGTGCAACC	CAACGGT--	-----	[402]
144_EFLCoelomomyces_stegomyiae	TTGGATTGGT	GATAACTTGA	AAGAAAAGTC	CAAG-----	-----AACA	TGGATTGG--	-----TGGAAAG	GGCCAAAAG	TATGCAACCT	TCTCAAG--	-----	[402]
145_EFLEntomophthora_muscae	ATGGATGGGT	GACAATCTTC	TCAAGAAGTC	AGAG-----	-----AAGA	TGACCTGG--	-----TGGAAAG	GGTGTGACG	TTAAGAACCA	GAGTGGC--	-----	[399]
146_EFLBasidiobolus_ranarum	ATGGCAAGGT	GATAACTTGT	TGACCAAGTC	TGCT-----	-----AAGA	TGCCCTGG--	-----TGGAAAG	GGTATGGATG	TTGTCAACCA	AGTTGGT--	-----	[399]
147_EFLConidiobolus_coronatus	CTGGCAAGGT	GATAACTTAC	TCAACCAATC	CACC-----	-----AACA	TGACCTGG--	-----TGGAAAG	GGTGTGATA	TTACCCGTAT	TGACGGT--	-----	[399]
148_EFLTetraselmis_tetrahelle	CTGGATGGGT	GACAACCTGC	TGTCCAAGTC	CCCC-----	-----AACA	TGGCTTGG--	-----TGGAAAG	GGTGTGTGAC	CTCTGGTCGG	CTCTGAG--	-----	[645]
149_EFLMesostigma_viride	GTGGGTGGGT	GACAACCTGC	TGAAGCCATC	TCCC-----	-----AAGA	TGGAGTGG--	-----TGGACG	GGTATGGATG	TCAAGAACCT	TGAGGGC--	-----	[699]
150_EFLOstreococcus_tauri	CTGGCAACGGT	GACAACCTGT	TGAAGAAGTC	CGAA-----	-----AAGA	TGGCTTGG--	-----TGGAAAC	GGTGTGACG	TTGAAGGCTA	CGACGGC--	-----A	[697]
151_EFLNephroselmis_olivacea	CTGGCTCGGC	GACAACCTGC	TGAAGCAGTC	CACC-----	-----AACA	TGGCTTGG--	-----TGGAAAG	GGTGCAGGAC	TCGACATCAA	CAAGAAG--	-----	[672]
152_EFLMicromonas_pusilla	CTGGCAGGGT	GACAACCTCA	TGCCCAAGAGC	CTCC-----	-----AACA	TGGGCTGG--	-----TGGTCC	GGCGTCGACG	TCATCGACTA	CGACCTC--	-----A	[646]
153_EFLParachlorella_kessleri	CTGGATGGGT	GACAATCTGA	TCAACCAAGTC	CACC-----	-----AACA	TGAGCTGG--	-----TGGACC	GGCGTGGATG	TGAACAACAT	GCTCAAC--	-----	[654]
154_EFLHelicosporidium_sp.	CTGGATGGGT	GACAACCTGA	TGAGAAGTC	CACC-----	-----AACA	TGAAGTGG--	-----TGGGAG	GGCGTCACCG	TGGACGTCT	GGACAAGGAC	GGCAAGAAGG	[709]
155_EFLGonium_pectorale	CTGGCTGGGC	GACAACCTGA	TCAACCAAGTC	GACC-----	-----AACA	TGCCCTGG--	-----TGACCT	GGCCAGACCG	TGGTCAACCT	GAAGAAC--	-----	[696]
156_EFLChlamydomonas_reinhardtii	CTGGCTGGGT	GACAACCTGA	TCAACCAAGTC	CACC-----	-----AACA	TGACCTGG--	-----TACTCC	GGCCAGGAGG	TGGTCAACCT	GAAGGGC--	-----	[696]
157_EFLScenedesmus_obliquus	CTGGATGGGT	GACAACCTGA	TCAAGCAGTC	CACC-----	-----AACA	TGGCTTGG--	-----TGGAAAG	GGCCAGGAGG	TGGTGTCTGC	TGACAAC--	-----	[696]
158_EFL"Oryza_sativa"	CTGGATGGGC	GACAACCTGA	TCAAGAAGTC	AGAG-----	-----AACA	TGAAGTGG--	-----TGGAAAC	GGCATGGAGG	TCAACCGGCT	GAACAAG--	-----	[696]
159_EFLChlorococcum_sp.	ATGGATGGGT	GACAACCTCA	TGCCCAAGTC	CGAG-----	-----AAGA	TGGGCTGG--	-----TGGTCT	GGATGAAGA	TCAAGTCCCT	CTCTGGC--	-----	[645]
160_EFLOchlochaete_hystrix	GTGGATGGGT	GACAACCTGA	TTGCCAAGTC	CACG-----	-----AACA	TGACCTGG--	-----TGGAAAC	GGCCAGGAGG	TGGACAACCT	TGAGGGC--	-----	[654]
161_EFLUlva_intestinalis	CTGGCATGGA	GACAACCTGA	TTGCCAAGTC	CAAA-----	-----AACA	TGACCTGG--	-----TGGAAAT	GGCCAGAAGG	TCAAGTCTCT	CGCTGGC--	-----	[645]
162_EFLUrospora_sp.	CTGGATCGGC	GACAACCTCA	TCAAGGCGAG	CGCC-----	-----AACA	TGACGTGG--	-----TGGAAAG	CCGACGGACA	TCAAGACGCT	CTCCGGC--	-----	[498]
163_EFLGymnochlorella_stellata_ppc	CTTCGGAGGC	GACAATTTAA	TTCAGAAGTC	GAAA-----	-----AATA	TGGATTGG--	-----TGGAAAC	GGATGTGATG	TCACTACCTC	ATGGGGT--	-----	[732]
164_EFLLotharella_amoebiformis_ppc	TTTCGAGGGC	GAGAACCTGC	TGAAGAAGTC	SGAA-----	-----AAAA	TGCCCTGG--	-----TGGAAAG	GGAAACGATG	TYGAAACCTC	GTGGGGC--	-----	[477]
165_EFLChlorarachnion_reptans_ppc	ATTCGAAAGGA	GACAACCTGA	TTGAGAAGAG	TAAA-----	-----AAGA	TGGATTGG--	-----TGGAAA	GGTAACGACA	TCGAGACTGT	ATACGGG--	-----	[504]
166_EFLChondrus_crispus	ATGGATGGGC	GATAACTTGA	TCACCNAGTC	TGCC-----	-----AAGA	TGGGTTGG--	-----TGGAAAG	GGCTGTGACG	TCCTTGTGCG	TAAGGAG--	-----	[696]
167_EFLGracilaria_changii	ATGGGAGGGT	GACAACCTGA	TACGAAAGTC	CACC-----	-----AACA	TGGGCTGG--	-----TGGAAAG	GGCTGTGATG	TGCTTGTGCG	AAAGAGG--	-----	[696]
168_EFLCryptomonas_ovata	CTGGATGGGC	GACAACCTCC	TCAAGAAGTC	TGAG-----	-----AACA	TGGCTTGG--	-----TGGAGT	GGCGTCGATG	TGAAGGCTGG	CAACGGC--	-----	[651]
169_EFLGuillardia_theta	ATGGATGGGA	GACAACCTCC	TCAACCAAGTC	CACC-----	-----AATA	TGGGATGG--	-----TGGTCC	GGAGTCGAGG	TGCTTCTGTA	TGGCTCC--	-----	[693]
170_EFLRhodomonas_salina	CTGGATGGGA	GACAACCTCC	TCAACCAAGTC	CACC-----	-----AACA	TGACCTGG--	-----TGGAAAC	GGTATCGAGA	TCGATGTCCC	CGACAAG--	-----	[693]
171_EFLPavlova_lutheri	CTGGATGGGC	GACAACCTGA	ACACCAAGTC	CGCC-----	-----AAGA	TGGACTGG--	-----TGGACC	GGCACGGACG	TCGACATCGC	GGCGGGC--	-----	[696]
172_EFLChrysochromulina_sp.	CTGGATGGGC	GACAACCTCA	TCAAGAAGAG	CGAG-----	-----AACA	TGTCGTGG--	-----TGGTCG	GGCCAGGACA	TCCTCGTCGA	GAAGGCC--	-----	[516]
173_EFLIsochrysis_galbana	ATGGATGGGC	GACAACCTCA	ATAAGAAGTC	GACC-----	-----AACA	TGGACTGG--	-----TACACG	GGACAGGACA	TCCTCTGCC	CAACAAG--	-----	[696]
174_EFLEmiliana_huxleyi	CTGGATGGGC	GACAACCTCA	TCGAGAAGAG	GAAG-----	-----AACA	TGGACTGG--	-----TGGACG	GGCTCGACG	TGATCGGGGG	CCCGGAC--	-----G	[697]
175_EFLAchnanthes_kuwaitensis	ATTCACCGGG	GACAACCTCG	TCGAGAAGTC	CGAC-----	-----AAGG	CTCTTTGG--	-----TACAAG	GGATGGACCG	CTAACATCAA	CCCCAAG--	-----	[615]
176_EFLAsterionella_glacialis	CTTCACCGGA	GATAACCTTG	TCACAAAGTC	CGAT-----	-----AAGG	CCCCCTGG--	-----TACAAG	GGATGGACCG	CCAACACTGC	CCCCAAG--	-----	[615]
177_EFLDetonula_confervacea	ATTTAACGGA	GACAACCTTG	TCGACAAGTC	CGAT-----	-----AAGG	CTCTTTGG--	-----TACAAG	GGCTGGACCG	CCAACATCAC	CCCCAAG--	-----	[615]
178_EFLThalassiosira_pseudonana	ATTTAACGGA	GACAACCTTG	TTGAGAAGAG	CGAT-----	-----AAGG	CTCTTTGG--	-----TACAAT	GGATGGGAGG	CCAACATCAA	CCCCAAG--	-----	[678]
179_EFLPerkinsus_marinus	TTGGATGGGC	GACAATCTGA	TCAAGAAGTC	CGAG-----	-----AAAA	TGTCTTGG--	-----TGGAAAG	GGTGTGGATG	TCACCTCGTA	CCCCGCT--	-----T	[706]
180_EFLOxyrrhis_marina	GTGGTGGCGT	GACAATTTGT	TCAAGAAGTC	CACC-----	-----AACA	TGCCGTGG--	-----TACAAG	GGACAGGATG	TTGTGGTTCC	GGGAACC--	-----	[702]
181_EFLKarlodinium_micrum	TTGGATGGGC	GACAACCTCT	TGAAGAAGTC	TGAA-----	-----AATA	TGGATTGG--	-----TGGAAAG	GGCTTCGACG	TCGCTGTTGG	CAAAGAG--	-----	[702]
182_EFLAmphidinium_carterae	CTGGATGGGC	GACAACCTTG	TCAAGAAGAG	TACC-----	-----AACA	TGGCTTGG--	-----TGGAGG	GGCGTCGATG	TTGAGGTTGG	AAAGGAG--	-----	[702]
183_EFLHeterocapsa_triquetra	CTGGATGGGT	GAGAAGAGTC	TGAAGAAGTC	GGAC-----	-----AACA	TGGCTTGG--	-----TGGAAAG	GGCCAGGACA	TCGAGTACGG	CAAGGAG--	-----	[702]
184_EFLPlanoglabratella_operculis	ATTTAAAGGA	GATAATTTAG	CAGAAGCCAG	CAAA-----	-----AATA	TGGCTTGG--	-----TACAAA	GGCTTTGAAA	TTAAAAAGAA	GGATAAG--	-----AATAAGA	[529]
185_EFLReticulomyxa_filosa	ATTTAAAGGA	GAAAACCTGA	CAAAAGTTAG	TGAA-----	-----AACA	TGCCATGG--	-----TGGAAAG	GGATTTTTCG	TCCAATTGAA	AAAGAT--	-----	[612]
186_EFLLotharella_amoebiformis	ATGGACTGGT	GACAACCTCA	TCCAGCCCTC	GACC-----	-----AAGA	TGCCCTGG--	-----TTCAAG	AGGTGGAAGG	CTACTGCCCC	CAACAAA--	-----	[765]
187_EFLGymnochlorella_stellata	ATGGTGTGGA	GACAACCTGA	TTATACCTTC	TACA-----	-----AAAA	TGCCTTGGTT	CAACAAGAAA	GGATGGACTG	CGACGACTCC	TTCGGA--	-----	[723]
188_EFLChlorarachnion_reptans	ATGGATGGGT	GATAACCTGA	TCAAGCCTTC	CACC-----	-----AACA	TGCCCTGGTT	TAACAAGAAA	GGGTGGACTG	CTACCAACCT	TAGCGGC--	-----	[654]
189_EFLDiplonema_ambulator	CTGGATCGGT	GACAACCTCC	TCAACCAAGTC	GGCG-----	-----AAGA	TGCCGTGG--	-----TGGAAAG	GGCCAGGACG	TCATCGTCGG	GGCGGAC--	-----	[297]
190_EFLDiplonema_papillatum	CTGGATCGGT	GACAACCTCC	TGAAGAAGTC	CGAG-----	-----AAGA	TGGCTTGG--	-----TGGAAAG	GGTGGCGACA	TCATCGTCGG	CAAGGAC--	-----	[669]
191_EFLRhynchopus_euleeides	CTGGATGGGT	GACAACCTCG	TGAAGAAGTC	CGAC-----	-----AAGA	TGGCTTGG--	-----TGGAAAG	GGCCAGGACG	TCRTCATCGA	GCAGGAC--	-----	[309]
192_EFLPetalomonas_cantuscygni	CTGGATGGGC	GACAACCTGC	TCAAGAAGTC	CGAG-----	-----AAGA	TGGCTTGG--	-----TGGAAAC	GGCCAGGACG	TCGCTGCCCA	GGACGGC--	-----	[540]
193_EFLTrypanoplasma_borrel	GTTCACGGC	GAGAACATGC	TGAAGAAGTC	CGAC-----	-----AAGA	TGCCGTGG--	-----TGGAAAC	GGCGTTGATG	TGAAGGTAGC	GGCCACC--	-----G	[715]
194_EFLNeobodo_salien	CTTCCACGGT	GAGAACCTGC	TCAAGGCGTC	CGAC-----	-----AAGA	TGCCGTGG--	-----TGGAAAG	GGGTTCACCG	TCTTCGACCC	GACGACC--	-----	[660]

	890	900	910	920	930	940	950	960	970	980	990
001_aEF1a_Desulfurococcus_mobilis	-----	-----	CCTACACTAG	TGGAGGCCACT	GGATACACTG	--GAGGTCC	CGCCTAAACC	CATAAACAAAG	CCGCTGAGGA	TACCGATACA	GGACGTCTAC
002_aEF1a_Sulfolobus_solfataricus	-----	-----	CCTACCTTAG	AAGAATATTT	AGATCAGTTA	---GAATTAC	CACCAAAACC	AGTAGATAAA	CCATTAAAGAA	TTCCAATTCA	AGACGTCTAC
003_aEF1a_Thermoplasma_acidophilum	-----	-----	cCAcAcAcTtC	tGcAgGcAcT	gGaTgCcTtC	---AaGgTaC	cTgAgAaGcC	caTaaAcAaG	cCgCtCaGgA	tCcCaGtTgA	aGaTgTgTaT
004_aEF1a_Halobacterium_sp.	-----	-----	CCCAACCTCC	TGGAGGCCCT	CAACGGGGTT	---CCGGTGC	CACAGCCGCC	GACGGACGCC	GATCTCCGCC	TCCCGATCCA	GGACGTCTAC
005_aEF1a_Methanocaldococcus_jannaschii	-----	-----	CCCACTCTCC	TTGAAGCATT	AGACAATTCT	---CAACCAC	CAGAAAAACC	ACCAAAATTA	CCATTAAAGAA	CCCAATCCA	AGATGTCTAT
006_EF1a_Acanthamoeba_culbertsoni	-----	-----	CCCAAGCTCC	TTGAGGCCCT	TGACGACATC	---AAGCCCC	CCAAGCGTCC	CATGGACAAG	CCCTCCCGTG	TGCCCTTCCA	GGACGTCTAC
007_EF1a_Neoparamoeba_branchipila	-----	-----	CCCACTCTCC	TGAGGCCCT	CGACTCCATC	---ATCCCCC	CCAAGCGTCC	CATCGACAAG	GCCTCCCGTG	TCCCCCTCCA	AGATGTCTAC
008_EF1a_Planoprotostelium_aurantium	-----	-----	CCCAACCTCC	TTGAGGCCCT	GGACTCCATC	---ACCGAGC	CCAAGAGACC	CTCAGACAAG	CCCTCCCGTG	TCCCCCTCCA	GGATGTGTAC
009_EF1a_Physarum_polycephalum	-----	-----	CCCAACCTCC	TGAGGCCCT	CGATCAGATC	---ACCGAGC	CCAAGCGCCC	GACCGACAAG	CCCTCCCGTG	TCCCCCTCCA	GGATGTGTAC
010_EF1a_Amoebidium_parasiticum	-----AA	CGCCACCAGC	AAGACCTGCT	TGGAGGCCCT	CGATGCCATC	---CTGCCGC	CCAAGCGCCC	CACGGACAAG	GCCTCCCGCC	TGCCCTTCCA	GGACGTCTAC
011_EF1a_Capsaspora_owczaraki	-----AA	CGCCTCCGGC	AAGACCTCTA	TTGAGGCTCT	TGACGCCATC	---TGCCCCC	CCAAGCGCCC	CACGGACAAG	CCCTCCCGCT	TCCCCCTCCA	AGACGTCTAC
012_EF1a_Monosiga_ovata	-----AA	TGCCAAAGGC	AAGACCTCCC	TGAGGCTCT	TGATGCCATC	---ATCCCCC	CGAGCGCCC	CACCTCCAAG	CCCTCCCGCC	TGCCCTTCCA	GGACGTCTAC
013_EF1a_Salpingoeca_amphoridium	CTGGTGGCAA	CGCCAAAGGC	AAGACTCTGC	TGGAAGCCCT	TGAGTGCATC	---CACCCCC	CTCAGCGCCC	CACCTGGCAA	CCCTCTCGCC	TGCCCTTCCA	AGACGTGTAC
014_EF1a_Ceanorhabditis_elegans	-----AA	CGCTTCCGGA	AAGACCTCCC	TGAGGCTCT	CGACTCCATC	---ATTTCCAC	CACAACGCCC	AACCTGACGA	CCACTCCGTC	TCCCCTTCCA	GGATGTGTAC
015_EF1a_Drosophila_melanogaster	-----AA	CGCTGACGGC	AAGACCTCGA	TGATGCCCT	CGATGCCATC	---CTTCCCC	CAGCCGCTCC	CACCGACAAG	GCCTTGCGTC	TGCCCTTCCA	GGATGTGTAC
016_EF1a_Ornithorhynchus_anatinus	-----AA	CGCCAGTGGG	GTATCCCTGC	TGGAAGCCCT	GGACCCATC	---CTGCCCT	CAGCCGCCCC	CACAGATAAG	CCCTTGCGTC	TTCCTCTCCA	AGATGTCTAC
017_EF1a_Xenopus_laevis	-----AA	TGCCAGTGGT	ATAACTCTGC	TAGAAGCACT	TGACTGTATA	---ATTCCCT	CTCAGAGACC	CACCTAACAA	CCCTTACGAC	TCCCTCTCCA	GGATGTGTAC
018_EF1a_Saccharomyces_cerevisiae	-----GT	CGTCAAGGGT	AAGACTTTGT	TGGAAGCCAT	TGACGCCATT	---GAACACA	CATCTAGACC	AACCTGACAAG	CCATTGAGAT	TGCCATTGCA	AGATGTTTAC
019_EF1a_Neurospora_crassa	-----AA	GGCCACTGGC	AAGACCTCCC	TGAGGCCAT	CGACGCCATT	---GAGCCCC	CCAAGCGTCC	TACCGACAAG	CCCTCCCGCT	TTCCTCTCCA	GGATGTCTAC
020_EF1a_Cryptococcus_neoformans	-----GT	TTCCAAGGGT	AAGACCTCCC	TGAGGCCAT	CGACGCCATC	---GAGCCCC	CTACCCGACC	CACCGACAAG	CCCTCCCGCT	TCCCTCTCCA	GGACGTCTAC
021_EF1a_Schizophyllum_comune	-----GT	CGTCAAGGGC	AAGACCTCCC	TGATGCCAT	CGACGCCATC	---GAGCCCC	CGCTTGTGCT	CTCCGACAAG	CCCTTCCGCT	TCCCCTTCCA	GGACGTCTAC
022_EF1a_Monoblepharis_insignis	-----AA	GGCCACTGGT	AAGACGCTCG	TTGAGGCCAT	TGACGCCATT	---GAGCCCC	CTAGCGCCC	CACCTGGCAA	CCCTCTCGCT	TCCCTCTCCA	GGATGTGTAC
023_EF1a_Rhizophlyctis_harderi	-----GC	TCTCAAGGGT	ACAACCTTCC	TCCAAGCTAT	TGATGCTATT	---GATCTCT	CAACCCGCTC	AACCTGACAA	CCCTTTCGTC	TCCCCTTCCA	AGATGTGTAC
024_EF1a_Harpochytrium_sp.	-----AA	GTTCACCGGA	AAGACCTCCC	TGAGGCCAT	CGACTCCATC	---GAGCCCC	CTTCGCGCCC	CACCGACAAG	CCCTTAGGTT	TGCCCTTCCA	GGATGTCTAC
025_EF1a_Absidia_glauca	-----GC	CAAGTCTGGC	AAGACTCTCT	TGGATGCCAT	CGATGCTATT	---GATCTCT	CCTAACGCTC	TCCGACAAG	CCCTTCCGCT	TCCCCTTCCA	AGATGTGTAC
026_EF1a_Smittium_simulii	-----GT	CTCTAAAGGT	GTCACTCTTC	TGATGCTAT	TGACGCTGTC	---GAACCTC	CAGTTGCTGC	AAGTGACAAA	CCCTTCCGCT	TTCTCTTCCA	AGACGTTTAC
027_EF1a_Chara_australis	-----	-----	CCCATTTCTC	TGACGCTCT	GGATCTGATT	---TCGGAGC	CGAAGAGGCC	GTTCTGACAAG	CCCTCTCGCT	TGCCCTTCCA	GGACGTTTAC
028_EF1a_Physcomitrella_patens	-----	-----	CGACTCTGT	TGGAGGCGTT	GGACAACGTC	---TCAGAGC	GGAAGCGGCC	ATCTGACAAG	CCCTTGGAGG	TGCCGCTTCA	GGACGTGTAC
029_EF1a_Acetabularia_acetabulum	-----	-----	CCCAACTTTT	TAGAGGCCCT	TGACATGGTG	---GAGCCCC	CCAACGCCCC	CACCGACAAG	CCCTTACGTT	TGCCCTTCCA	AGATGTCTAC
030_EF1a_Chaetomorpha_coliformis	-----	-----	CCCAAGCTCT	TAGAGGCCCT	GGACAATGTC	---GACCCCC	CAAGAGAGCC	CTCCGACAAG	CCCTTACGTC	TCCCCTTCCA	GGATGTGTAC
031_EF1a_Arabidopsis_thaliana	-----	-----	CCCACTCTTC	TTGAGGCTCT	TGACAGATC	---AACGAGC	CAAGAGAGCC	ATCAGACAAG	CCCTTTCGTC	TCCCCTTCCA	GGATGTCTAC
032_EF1a_Oryza_sativa	-----	-----	CCCACTCTTC	TGAGGCTCT	TGACAGATC	---AACGAGC	CCAAGAGGCC	ATCAGACAAG	CCCTTACGTC	TCCCCTTCCA	GGACGTGTAC
033_EF1a_Cyanidioschyzon_merolae	-----	-----	CCAGCCCTGC	TGGAGGCCCT	TGATGCCATC	---GAGCCTC	CGAAGCGCCC	CACCGACAAG	CCACTGCGTC	TGCCCTTCCA	GGATGTGTAC
034_EF1a_Porphyrha_jezoensis	-----	-----	CGTGCCCTGC	TGGAGGCTCT	GGACAACCTG	---GACCCGC	CGAAGCGCCC	GGTTGACAAG	CCGCTGCGCC	TGCCCTTCCA	GGACGTCTAC
035_EF1a_Plasmodium_falciparum	-----	-----	AGAACCTTAA	TAGAAGCCCT	TGATACTATG	---GAACACC	CAAGAGAGCC	TTACGACAAG	CCATTAAAGAA	TTCCACTTCA	AGGTGTTTAC
036_EF1a_Cryptosporidium_parvum	-----	-----	AAGACCTTAG	TGCAAGCCCT	CGACACTATG	---GAACACC	CAAGAGAGCC	AACCTGACAAG	CCACTCCGTC	TCCCATTACA	AGATGTTTAC
037_EF1a_Tetrahymena_pyriformis	-----	-----	CCCATCTCTC	TGCAAGCTCT	CGATGCTCTT	---GAACCCC	CAAAAGAGCC	TGTTGACAAG	CCCTTACGAC	TCCCCTTCCA	AGACGTCTAC
038_EF1a_Stentor_coeruleus	-----	-----	CCAACCTCTC	TAGGAGCTCT	TGATATGATC	---ATCCCTC	CTAAGCGTCA	CTCTGACAAG	CCGTTGCGCT	TCCCCTTCCA	GGACGTTTAT
039_EF1a_Phytophthora_parasitica	-----	-----	CCCATCTCTC	TGAGGCTCT	TGACAACCTG	---AACCCCC	CTAAGCGCCC	GGTTGACAAG	CCGCTGCGCT	TCCCTTCCA	GGATGTGTAC
040_EF1a_Mallomonas_rasilis	-----	-----	CCCTTCTCTC	TGAGGCTCT	TGACTTAGTG	---AACCCCC	CTAAGAGACC	CACCTGATAAG	CCCTTTCGTC	TCCCCTTCCA	GGATGTATAC
041_EF1a_Phaeodactylum_tricornutum	-----	-----	CCTTACTCTG	TGAGGCCCT	TGACAGTGTA	---ACCCCGC	CCAAGCGTCC	CACCGACAAG	GCTCTACGAC	TCCCCTTCCA	GGATGTCTAC
042_EF1a_Thalassiosira_pseudonana	-----	-----	CTTACCCTTA	TGCAAGCATT	GGACAACGTT	---CACCTCT	CCAAGCGTCC	TATTGACAAA	CCCTTTCGGC	TGCCCTTCCA	AGATGTGTAC
043_EF1a_Malawimonas_jakobiformis	-----	-----	TTCACTCTGT	TGGAGGCCCT	GGACAACCTG	---GAGCCCC	CGAAGCGCCC	GTTTGAAGAA	CCGCTGCGCT	TGCCCTTCCA	GGACGTGTAC
044_EF1a_Malawimonas_californiana	-----	-----	CCCACGCTGC	TGGAGGCCCT	CGACGCCATC	---GAGCCCC	CCAAGCGCCC	CTCCGAGAAG	CCCTTCCGCT	TGCCCTTCCA	GGACGTCTAC
045_EF1a_Trimastix_pyriformis	-----	-----	CCCACTCTTC	TTGGTGCTCT	GAACAACCTC	---GAGGTCC	CCAAGCGCCC	CGTTCACAAG	CCCTTCCGCT	TCCCCTTCCA	GGATGTCTTC
046_EF1a_Dinenympha_exilis	-----	-----	CCAATTTTGT	TGATGCTCT	GGATCTTTTG	---GAAGTGC	CGAAGCGTCC	CACCGACAAG	CCCTTCCGCT	TCCCCTTCCA	GGATGTGTTC
047_EF1a_Monocercomonoides_sp.	-----	-----	GATATTTCTC	TTGATGCTCT	TGACAACCTG	---GAAGTTC	CAAGAGAGCC	AATTGACAAG	CCCTTCCGCT	TCCCCTTCCA	GGATGTGTTC
048_EF1a_Streblospio_saxatilis	-----	-----	AACACATTAT	TGATTTCACT	TGACACACTC	---GAAGTTC	CAAAACGCTC	AGTTGACAAA	CCATTGAGAC	TTCGATTACA	GGATGTATT
049_EF1a_Histomonas_melagris	-----	-----	GGCAACACTC	TTGAAACACT	CGATACACTC	---CAACACC	CAAAAGCGTCC	ATTGACAGCT	CCACTCCGTC	TTCGATTACA	AGATGTCTAC
050_EF1a_Trichomonas_vaginalis	-----	-----	CTACACTCTC	TTGAAGCCCT	CGATTTCCCTT	---CAGCCAC	CAAAAGCGCCC	ATTGACACAAG	CCACTCCGTC	TTCGATTACA	GGATGTCTAC
051_EF1a_Hexamita_inflata	-----	-----	AAGTGCCCTA	TGCAATGCAT	CGATGGTCTC	---AAGGCTC	CAAAAGCGTCC	AAACGACAAG	CCACTCCGCT	TCCCATTACA	AGACGTCTAC
052_EF1a_Spiroplasma_vortens	-----	-----	CCGTGCCCTA	TGAGGCCAT	CGACTCCCTC	---AAGCCCC	CCAAGCGCCC	CACCTGACAAG	TGCTTCCGCT	TCCCCTTCCA	GGACGTCTAC
053_EF1a_Andalucia_incarcerata	-----	-----	CCCACTCTC	TTGAGGCTCT	CGATGGTCTT	---GAGGAGC	CTTCTCGCCC	CGTTCACAAG	CCCTTCCGCT	TCCCCTTCCA	GGATGTTTAC
054_EF1a_Seculamonas_ecuadoriensis	-----	-----	CCGACCCCTG	TGGAGGCCAT	CGACTCGTCT	---GAGGAGC	CGAAGCGCCC	GTCGACAAG	CCGCTGCGCT	TGCCCTTCCA	GGACGTCTAC
055_EF1a_Reclinomonas_americana	-----	-----	CCGATCTTGA	TTGAGGCCAT	TGACCACTTC	---GAGGAGC	CGAAGCGCCC	GACGACAAG	CCCTTCCGCA	TCCCCTTCCA	GGACGTGTAC
056_EF1a_Histiona_aroidea	-----	-----	CAACCCCTGC	TGAGGCCCT	GGATCAATTC	---GAAGAGC	CAAAAGCGCCC	CACCGACAAG	CCACTCCGCT	TGCCATTGCA	AGATGTGTAC
057_EF1a_Jakoba_libera	-----	-----	CCCAACCTTC	TGGAGGCCAT	CGACACTCTC	---AGCGAGC	CCAAGCGCCC	CCACGACAAG	CCCTTCCGCT	TCCCCTTCCA	GGACGTGTAC
058_EF1a_Stachyamoeba_lipophora	-----	-----	CCCACTCTGT	GGAGGACGCT	TGACCTCCTC	---GAGGCCC	CCAAGCGTCC	CACCGACAAG	CCCTCTCGTA	TCCCCTTCCA	GGATGTGTAC
059_EF1a_Naegleria_gruberi	-----	-----	CCATGTCTTT	TGATGCTTTT	GGACAACCTG	---ATTGAAC	CAGTTAGACC	AACCTGAAAAG	CCATTGAGAT	TGCCACTTCA	AGACGTTTAC
060_EF1a_Naegleria_andersoni	-----	-----	CGATGTTTGT	TGATGCTTTT	GGATAACTTG	---GTTGAAC	CTGTTAGACC	AACCTGATAAG	CCATTGAGAT	TGCCACTTCA	AGACGTATAC
061_EF1a_Acrasis_rosea	-----	-----	CCCACTCTAT	TGGAAGCCCT	GGACGCTCTT	---GAGCCAC	CAAGAGAGCC	AACGAGAAG	CCACTACGTC	TACCACTACA	AGATGTCTAC
062_EF1a_Sawyeria_marylandensis	-----	-----	CCAACATTAT	TAGAAGCATT	AGATATGATT	---AAACCAC	CAGAGAAACC	AACAGATAAA	CCATTAAAGAT	TACCACTACA	AGATGTATAC
063_EF1a_Bodo_saltans	-----	-----	CCCAACCTGC	TGGAGTCCCT	CGACTCGCTC	---GAGCCCC	CGACCCGCCC	CTCGGACAAG	CCCTTCCGCC	TGCCCTTCCA	GGACGTGTAC
064_EF1a_Rhynchobodo_sp.	-----	-----	CCCAACCTGC	TGAGGCCCT	CGACGCCCTC	---GAGCCCC	CGCGCGCCC	AAGCGACAAG	CCCTTCCGCC	TGCCCTTCCA	GGACGTCTAC
065_EF1a_Rhynchomonas_nasuta	-----	-----	CCCAACCTGC	TGAGGCCCT	CGACATGCTC	---GAGCGCC	CGCTGCGCCC	GTCGACAAG	CCGCTGCGCC	TGCCCTTCCA	GGACGTCTAC
066_EF1a_Dimastigella_trypaniformis	-----	-----	CCGACCCCTG	TGGAGGCCCT	GGACTCCCTG	---GAGGCCC	CGACCCGCCC	GACGACAAG	CCGCTGCGCC	TGCCCTTCCA	GGACGTCTAC

	890	900	910	920	930	940	950	960	970	980	990	
067_EF1a_Perkinsiella_amoebae	-----	-----	CCAACCTCC	TCGAGGCCCT	CGACAGCTC	---GACCCG	CCGTGCGCCC	CGTTGACAAG	CCGCTCCGCC	TGCCGCTCCA	GGACGTCTAC	[651]
068_EF1a_Leishmania_major	-----	-----	CCACAGCTGC	TGGACGCGCT	CGACATGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	CGCTGCTGAC	GGACGTGTAC	[726]	
069_EF1a_Leishmania_infantum	-----	-----	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGACGTGTAC	[726]	
070_EF1a_Leishmania_braziliensis	-----	-----	CCCACGCTGC	TGGACGCGCT	GGACATGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGACGTGTAC	[726]	
071_EF1a_Leishmania_tarentolae	-----	-----	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGACGTGTAC	[726]	
072_EF1a "L. tarentolae" FJ807257	-----	-----	CCGACGCTGC	TGGAGGCGCT	GGACATGCTG	---GACCAC	GGTGGACAAG	CCGCTGCGCT	TGCCCTGCA	GGACGTGTAC	[651]	
073_EF1a_Leptomonas_costaricensis	-----	-----	CCGACGCTGC	TGGACGCGCT	CGACATGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGACGTGTAC	[474]	
074_EF1a_Leptomonas_bifurcata	-----	-----	GCCACGACGC	TGGAGGCTCT	TGACCTGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGATGTGTAC	[486]	
075_EF1a_Leptomonas_podlipaevi	-----	-----	CCCACGACGC	TGGAGTCCCT	CGACATGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGATGTGTAC	[474]	
076_EF1a_Herpetomonas_muscarum	-----	-----	CCCACGCTGC	TGGAGGCGCT	CGACAGCTG	---GACCCC	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGACGTGTAC	[639]	
077_EF1a_Herpetomonas_pessoai	-----	-----	CCCACCTTGC	TGGAGGCGCT	CGACTCCCTG	---GACCCC	GGTGGACAAG	CCCTGCGCC	TGCCCTGCA	GGACGTGTAC	[597]	
078_EF1a_Blastocrithidia_culicis	-----	-----	ATGACGCTGC	TGGAGTGCCT	CGACTCCCTG	---GAGCCG	GGTGGACAAG	CCCTGCGCC	TGCCCTGCA	GGACGTGTAC	[639]	
079_EF1a_Trypanosoma_brucei	-----	-----	CCAACGCTCC	TGGAGGCACT	AGACATGCTG	---GAGCCG	GGTGGACAAG	CCCTGCGCT	TGCCCTGCA	GGACGTGTAC	[726]	
080_EF1a_T. brucei_equiperdum	-----	-----	CCAACGCTCC	TGGAGGCACT	AGACATGCTG	---GAGCCG	GGTGGACAAG	CCCTGCGCT	TGCCCTGCA	GGACGTGTAC	[648]	
081_EF1a_Trypanosoma_aviium	-----	-----	CCCACCTTGC	TGGAGGCACT	CGATATGCTG	---GAGCCG	GGTGGACAAG	CCCTGCGCT	TGCCCTGCA	GGACGTGTAC	[672]	
082_EF1a_Trypanosoma_cruzi	-----	-----	CCCACGCTGC	TGGAGGCACT	CGACATGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGATGTGTAC	[726]	
083_EF1a_Trypanosoma_rangeli	-----	-----	CCCACCTTGC	TGGAGGCACT	CGACATGCTG	---GAGCCG	GGTGGACAAG	CCGCTGCGCC	TGCCCTGCA	GGACGTGTAC	[726]	
084_EF1a_Ploetitia_costatum	-----	-----	AACACCTTGA	TCCAGGCTCT	TGATGCTTTG	---GAGCCG	CCAAGCGACC	CAATGACAAG	CCCTTCCGAC	TGCCCTTCCA	GGACGTGTAC	[723]
085_EF1a_Ploetitia_oblonga_cDNA	-----	-----	AACACCTTGA	TCCAGGCTCT	TGATGCTTTG	---GAGCCG	CCAAGCGACC	CAATGACAAG	CCCTTCCGAC	TGCCCTTCCA	GGACGTGTAC	[723]
086_EF1a_Ploetitia_oblonga_gDNA	-----	-----	CCCACGCTCA	TGGAGGCGCT	CGACAACTTG	---GAGCCG	CCAAGCGCCC	CGCCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGACGTGTAC	[723]
087_EF1a_Entosiphon_sp.	-----	-----	CCGTGCTTCA	TGGAGGCGCT	CGACGCTTTG	---GATGCC	CCAAGCGCCC	CGTTGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGACGTGTAC	[723]
088_EF1a_Entosiphon_sulcatum	-----	-----	CCGTGCTTCA	TGGAGGCGCT	CGACGCTTCT	---GAGCCG	CCAAGCGCCC	CTCCGATAAG	CCCTTCCGCT	TGCCCTTCCA	GGACGTGTAC	[651]
089_EF1a_Peranema_tricophorum	-----	-----	TCAACATTTGA	TGATGCTTCT	CGACAAATTA	---GAACCA	CTAAAGGACC	AAACGACAAG	CCACTCAGAC	TGCCCTTCCA	GGATGTGTAC	[723]
090_EF1a_P. tricophorum_FJ807243	-----	-----	TCAACCTTGA	TGATGCTTCT	CGACAAATTA	---GAACCA	CTAAAGGACC	AAACGACAAG	CCACTCAGAC	TGCCCTTCCA	GGATGTGTAC	[717]
091_EF1a_Colacium_elongatum	-----	-----	CCCACCTTCA	TGGGCGCTCT	GGACAACCTTG	---GAGCCG	CCAAGCGTCC	CAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
092_EF1a_Colacium_mucronatum	-----	-----	CCGACCTTGA	TGGGCGCTTT	GGATAACTTG	---GAGCCG	CCAAGCGGCC	CAACGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[669]
093_EF1a_Colacium_sp.	-----	-----	CCCACCTTCA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CCAAGCGGCC	CAACGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
094_EF1a_Colacium_vesiculosum	-----	-----	CCGACCTTGA	TGGGCGCTTT	GGATAACTTG	---GAGCCG	CCAAGCGGCC	CAACGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[681]
095_EF1a_Cryptoglena_pigra	-----	-----	CTGACTCTGA	TGGGCGCTTT	GGAGCGCTTG	---GAGCCG	CCAAGCGTCC	CAGCGACAAG	CCGCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
096_EF1a_Cryptoglena_schubertii	-----	-----	CTGACTCTGA	TGGGCGCTTT	GGAGCGCTTG	---GAGCCG	CCAAGCGTCC	CAGCGACAAG	CCGCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
097_EF1a_Discoplistis_adunca	-----	-----	CCCACGCTCA	TGGGCGCTTT	GGACATGCTG	---GAGCCG	CCAAGCGTCC	CAGCGACAAG	CCGCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
098_EF1a_Euglena_clara	-----	-----	CTGACCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CCAAGCGTCC	CAGCGACAAG	CCGCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[693]
099_EF1a_Euglena_deses	-----	-----	TTGACCTTGA	TGGGCGCTTT	TGACAACCTTG	---GAGCCG	CWAAAGCGCC	AAGTGACAAG	CCACTTCCGAC	TGCCCTTCCA	GGATGTGTAC	[723]
100_EF1a_Euglena_exilis	-----	-----	CTTACTTCTGA	TGGGCGCTTT	GGACAACCTTG	---GAACCA	CCAAGCGGCC	AAATGACAAG	CCCTTCCGAC	TGCCCTTCCA	GGATGTGTAC	[693]
101_EF1a_Euglena_geniculata	-----	-----	AACACCTTGA	TCCAGGCTCT	TGATGCTTTG	---GAGCCG	CCAAGCGACC	CAATGACAAG	CCCTTCCGAC	TGCCCTTCCA	GGATGTGTAC	[723]
102_EF1a_Euglena_gracilis	-----	-----	CTGACCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CCAAGCGACC	AAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
103_EF1a_E. gracilis_X16890	-----	-----	TTGACCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAACCG	CAAAGCGTCC	TGCCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[726]
104_EF1a_E. gracilis_FJ807238	-----	-----	CTGACCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGCGTCC	AAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[726]
105_EF1a_Euglena_granulata	-----	-----	AATACTTCTGA	TCCAGGCGCT	GGACTCCTTG	---GAGCCG	CTAAAGGACC	TAATGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
106_EF1a_Euglena_hiemalis	-----	-----	CTCACCTTGA	TCCAGGCGCT	GGACAACCTTG	---GAGCCG	CAAAGCGGCC	CAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
107_EF1a_Euglena_longa	-----	-----	CTGAGCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGCGGCC	AAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[240]
108_EF1a_Euglena_stellata	-----	-----	TTGACACTTA	TGGGCGCTTT	GGACAACCTTG	---GAAGCA	CAAAGCGGCC	TTCTGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[693]
109_EF1a_Euglena_viridis	-----	-----	CTGACTTCTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CTAAAGGACC	ATCTGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[681]
110_EF1a_Eutreptia_pertyi	-----	-----	AATACTTCTGA	TCCAGGCGCT	GGACTCCTTG	---GAGCCG	CAAAGCGTCC	TAATGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
111_EF1a_Eutreptiella_pomquetensis	-----	-----	AACACCTTGA	TCCAGGCTCT	TGATGCTTTG	---GAGCCG	CCAAGCGACC	CAATGACAAG	CCCTTCCGAC	TGCCCTTCCA	GGATGTGTAC	[333]
112_EF1a_Lepocinlis_acus_var._major	-----	-----	ATGACCTTGA	TCCAGGCGCT	GGATGCTTTG	---GAGCCG	CTAAAGGACC	TACCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
113_EF1a_Lepocinlis_ovum	-----	-----	ATGACCTTGA	TCCAGGCGCT	GGAGCGCACT	---GAGCCG	CAAAGCGTCC	TACCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[693]
114_EF1a_Lepocinlis_spyrogyroides_clone_38	-----	-----	-----	-----	-----	---	-----	-----	-----	-----	-----	[36]
115_EF1a_L. spyrogyroides_clones_1_&_4	-----	-----	CTGACTTCTGA	TCCAGGCGCT	TGACAACCTTG	---GAGCCG	CTAAAGGACC	CAGCGACAAG	CCACTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
116_EF1a_Lepocinlis_steinii	-----	-----	CCGACGCTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGCGACC	ATCCGACAAG	CCACTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[693]
117_EF1a_Lepocinlis_triptyeris	-----	-----	CTGACTTCTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGGCGCC	CTCTGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
118_EF1a_Monomorphina_aenigmatica	-----	-----	CTGACTTCTGA	TGGGCGCTTT	GGATGCTTTG	---GAGCCG	CAAAGGCGTCC	CAATGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
119_EF1a_Monomorphina_pyrum	-----	-----	CTGACTTCTGA	TGGGCGCTTT	TGACAACCTTG	---GAGCCG	CTAAAGGACC	CAGCGACAAG	CCACTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
120_EF1a_Monomorphina_sp.	-----	-----	CTGACTTCTGA	TGGGCGCTTT	GGACAACCTTG	---GAACCG	CAAAGGCGCC	CAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
121_EF1a_Phacus_orbicularis	-----	-----	AACACTTCTGA	TGGAGGCTCT	GGACAACCTTG	---GAGCCG	CAAAGGCGTCC	CTCTGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[681]
122_EF1a_Phacus_similis	-----	-----	CTGAGCTTGA	TGGGCGCTTT	AGATAACTTG	---GAGCCG	CAAAGGCGCC	GTGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
123_EF1a_Strombomonas_acuminata	-----	-----	CTGACTTCTGA	TGGGCGCTTT	GGATAACTTG	---GAACCT	CAAAGGCGTCC	TTCTGACAAG	CCACTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[726]
124_EF1a_Strombomonas_borsteniensis	-----	-----	CTGACTTCTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGGCGCC	GAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[693]
125_EF1a_Strombomonas_napiformis	-----	-----	CTCACCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGGCGCC	AAATGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
126_EF1a_Strombomonas_ovalis	-----	-----	CTGACTTCTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGGCGCC	AAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
127_EF1a_Strombomonas_triquetra	-----	-----	AACACCTTGA	TGGGCGCTTT	GGATGCTTTG	---GAGCCG	CAAAGGCGCC	CAGCGACAAG	CCACTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
128_EF1a_Trachelomonas_conradi	-----	-----	CTGACCTTGA	TGGGCGCTTT	GGACAACCTTG	---AAGCCG	CAAAGGCGCC	CAGCGACAAG	CCACTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
129_EF1a_Trachelomonas_echinata	-----	-----	TTGACACTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGGCGCC	AAATGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
130_EF1a_Trachelomonas_lefevrei	-----	-----	CTGACCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGGCGTCC	AAGCGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[723]
131_EF1a_Trachelomonas_planctonica	-----	-----	CTGACCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGGCGCC	AAGTGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[693]
132_EF1a_Trachelomonas_similis	-----	-----	TTGACCTTGA	TGGGCGCTTT	GGACAACCTTG	---GAGCCG	CAAAGGCGCC	AAGTGACAAG	CCCTTCCGCT	TGCCCTTCCA	GGATGTGTAC	[693]

	890	900	910	920	930	940	950	960	970	980	990	
133_EFL1a_Trachelomonas_volvocina	-----	-----	CACACCCCTCA	TTGGTGTCT	GGACAACCTTG	---GAGCCTC	CCAAGCGCTCC	AAGTGACAAG	CCTCTGCGTG	TTCTCTGCGA	GGATGTCTAC	[723]
134_EFL_environmental_sample_EU810232	-----AC	CGTGCACGTG	CACACCTTGC	TCGACGCTCT	TCGAGAAGATG	GTCAACCCCGC	CCGAGACGGTC	CGACAAGGGA	GCCCTCCGTG	TGCCCATCAG	CGGAGTGTAC	[798]
135_EFL_Creolimax_fragrantissima	-----AA	GATCCACGTA	CACACTCTTT	TGGATGCGCT	TGAGTTCATG	GTCAACCATTC	CCACCCGTAC	CCCAGATCTT	GCCATGCGTA	TGCCCTATCT	TGGAGCCTTC	[726]
136_EFL_Sphaeroforma_arctica	-----AA	GGTTACAGTT	GACTGTCTCT	TTGACTGTCT	TGAGTTCATG	GTTCACCATCC	CGGTGCGTAC	CCCCGATCTA	ACCATGCGTA	TGCCCTATCT	AGGAGCCTTC	[798]
137_EFL_Monosiga_brevicollis	-----AC	CATCCACATT	GACTGCCCTCC	AGGATGCTCT	CGAGAAGATG	GTCCGCTTCC	CCAGCGCGCC	ACCAGACAAG	CCCATGCGTA	CCCCATCTCT	GGGTGTCTTC	[798]
138_EFL_Xenopus_laevis	-----AA	GATCCACATC	GACACCCGTG	TCGATGCCCT	GGAGAAGATG	GTGACAGTGC	CCGCCCGCCC	CGTGACAGCC	AAGCTGCGTG	TGCCCTGTC	TGGCGCCTTC	[798]
139_EFL_Spizellomyces_punctatus	-----AAGAA	GGTCAAGGTA	CACACTCTCC	TCGATGCCCT	GAACGACTTC	GCTGAGATGC	CGAGCGGTAA	GAACGACGCG	CCCATGCGTG	TCCCCATCTC	TGGTATCTAC	[804]
140_EFL_Gaertneriomyces_semiglobiferous	-----AAGAA	GATCCAGGTC	CACACCCCTC	TCGACGCCCT	CAACGACTTC	GCCGAGATGC	CGAGCGGTAA	GAACGACGCC	CCCATGCGTG	TCCCATCTCT	TGGTATCTA-	[506]
141_EFL_Allomyces_macrogynus	-----AAGGA	CATCATGATC	GTGACGCTGC	TCAACGCCCT	CAACGACTTC	GCTGAGCTGC	CGCCGCGCAA	GTGACGCGCC	CGCATGCGCG	TGCCCATCTC	GGGTATCTAC	[804]
142_EFL_Blastocladiella_emersonii	-----AAGAA	GGTCAAGATC	GTGACCTCC	TCAACGCCCT	CAACGACTTC	GCTGAGCTCC	CCGCTCGCAA	GTGCGGCTCG	CCCTCCGTC	TCCCCATCTC	GGGATCTAC	[804]
143_EFL_Catenophylctis_sp.	-----TCCAA	GCACACCATT	GTACCCCTCC	TCAACGCCCT	CAACGACTTC	GCCGATCTCC	CCACCCGCAA	GTGCGACGCT	GCCCTCCGTC	TCCCCATCTC	GGGTATCTA-	[506]
144_EFL_Coelomomyces_stegomyiae	-----AAGA	AGTCCAGGTT	GACACCTTAG	TAGATGCTTT	GAATACTTTT	GCTGAATTAC	CCCCAGAA	GAGCGATGCT	GCCATGCGTG	TCCCCATTTT	CGGTATTTA-	[506]
145_EFL_Entomophthora_muscae	-----AAGGA	TTTCCATATC	GAGACCCCTC	AGGTTGCTCT	TAACGACTTC	GTGAGCTTC	CCTCCGCTGT	CGATGACGCT	CCCATGCGTG	TCCCCGTTTC	CGGTATCTA-	[503]
146_EFL_Basidiobolus_ranarum	-----GACAA	GATCCACATT	GAYACTCTTC	TCMACGCTTT	GAACGAGATG	GTGTCCTCC	CCACTCGTAA	CACCGAAGCT	GCTATGCGTG	TYCCCGTTTC	CGGTATCTA-	[503]
147_EFL_Conidiobolus_coronatus	-----AAGAA	GATCCACGTC	CATACTCTTA	AGGATGCTTT	AAACGATATG	GTTCACATTC	CCCAACGTAA	CGTGCATGCC	CCATGCGTG	TCCCCGTTTC	CGGTATCTA-	[503]
148_EFL_Tetraselmis_tetrathele	-----AA	GCTGCACATC	GAGTGCCCTGA	AGGATGCCCT	CAACGACTTC	TTCCCGCTAC	CCAGCGCAA	CAACACCGCC	AAGCTGCGCG	TGCCCATCTC	CGGTATCTAC	[747]
149_EFL_Mesostigma_viride	-----AAGAC	TATCAAATGC	CATACCTTCC	TCAACGTGCT	CAACGACTAC	GCCACCGTGC	CCGAGCGCAA	GACCGATGCC	CCCATGCGCC	TGCCCATCTC	CGGTATCTAC	[804]
150_EFL_Ostreococcus_tauri	AGGGCGCCAA	GTTCACATT	GACACCCCTC	TCAACGTCTC	CAACGACTGC	GTGACGCGCC	CGGACCGCAA	CACCGAAGCC	CAAAATGCGCC	TCCCGATCTC	CGGTGTCTAC	[807]
151_EFL_Nephrolepis_olivacea	-----AA	GATGAAGGTT	ACACCCCTCC	TGGACTGCCCT	GAACGACATG	GTTCGAGCTC	CGCGCCGCGA	CACCGCCTCC	CCCATGCGCC	TGCCCGTGAA	CGGATCTCTAC	[774]
152_EFL_Micromonas_pusilla	AGGGCACCAC	GCACACATC	GACACCCCTC	TCGATGTCTC	CAACRCGATG	GTCAACCAAGC	CGCCCGCAA	CGACAAGGCT	GACATGCGCC	TGCCCATCTC	YGGCGTGTAC	[756]
153_EFL_Parachlorella_kessleri	-----AAGAA	GATCCACGTC	CACACCCCTG	TGGACGCACT	CAACGACTTT	GCTGAGGTGC	CGAGCGCAA	GACCGACGCG	CCCTGCGCC	TGCCCATCTC	TGGTATCTAC	[759]
154_EFL_Helicospiridium_sp.	TGGAGGAGAA	GGTCAAGGTC	CACACCCCTG	TGGACGCCCT	GAACGACATG	GTTCGCCCTG	CCGAGCGCAA	GATCGACGCC	CCCGTCCGCG	TGCCCATCTC	GGGATCTCTAC	[819]
155_EFL_Gonium_pectorale	-----GAGAA	GATCCAGGTC	CACACCCCTC	TGGACGCCCT	GAACGACTTC	GTCACCGTGC	CGAGCGCAA	GGTCAACGCG	CCCTGCGCC	TGCCCATCTC	GGGATCTCTAC	[801]
156_EFL_Chlamydomonas_reinhardtii	-----GAGAA	GATCCAGGTC	CACACCCCTG	TGGACGCCCT	GAACGACTTC	GTGGTGTGTC	CGAGCGCAA	GACCGACGCC	CCCTGCGTC	TGCCCATCTC	CGGCGCTCTAC	[801]
157_EFL_Scenedesmus_obliquus	-----CGCAA	GGTGCAGATC	AAGTGCCTGC	TGGACGCCCT	GAACGACTGC	GTGATTGTGC	CTGAGCGCAA	GACCGACGCC	GCCATGCGTG	TGCCCATCTC	CGGCGCTCTAC	[801]
158_EFL_Orzya_sativa	-----GAGAA	GGTGCAGATC	GAAGTGCCTG	TGAATGCCCT	GAACGACTGC	GTGATGTGTC	CTGAGCGCAA	GACCGACGCC	CCCATGCGCG	TGCCCATCTC	CGGCGCTCTAC	[801]
159_EFL_Chlorococcum_sp.	-----AAGGA	GGCCAGGTC	AACACCCCTC	TGGACTACCT	CAATGACGTT	GTGAGGTGTC	CTGAGCGCAA	GACTGATGCC	GCCATGCGCG	TGCCCATCTC	TGGTATCTAC	[750]
160_EFL_Ochlochaete_hystrix	-----AAGAA	GATCCAGGTC	ACCACTGTGC	TGGACGCCCT	GAACGACTTT	GCTGTGATGC	CGAGCGCAA	GGTGAATGCC	CCCATGCGCG	TGCCCATCTC	CGGTATCTAC	[759]
161_EFL_Ulva_intestinalis	-----AAGGA	GGTCAAGTGC	GTGACACTCC	TCAACTACCT	GAACGAGGTC	GTGAGGTGTC	CGCTCGTGA	CACCGACAAG	CCCTCCGTC	TGCCCATCTC	CGGTATCTAC	[750]
162_EFL_Urospora_sp.	-----AAGCC	GGTGCAGGTC	GAGACCCCTG	TGGACTACCT	CAACGTGTC	CGAGGCTGTC	CGAGCGCAA	CACGAGAAG	CGCGTGC	TGCCCATCTC	GGGATCTCTAC	[603]
163_EFL_Gymnochlorella_stellata_ppc	-----GATAG	CATACATGTC	AATACACTGC	TTGACTGCTT	AAACGATATG	GTTCGAGCTC	CGCCCGGAAA	AAGGATGCTT	CCCTTGC	TGCCCATCTC	TGGTGTGATC	[837]
164_EFL_Lotharella_amoebiformis_ppc	-----GARAA	GATTACGTCG	GATACTCTTC	TAGACTGCTT	GAACGACTGC	GTTCAGACCGC	GTCCCGGAAA	GGTGAACGCT	CCCTGCGCTG	TGCCCATCTC	CGGCGTGTAC	[582]
165_EFL_Chlorarachnion_reptans_ppc	-----AAAAA	ACTACACGTC	GATACCCCTG	ATGACTGCTT	CGATGAAATG	GCACAGCCCC	CAGAGAGAAA	AGTGCATGCT	CGGTGCGCG	TTCTATCTCG	CGGCGTTTAC	[609]
166_EFL_Chondrus_crispus	-----AA	GGCCATATCT	GACACTTTGT	CCGATGCTCT	CGAGAAGATG	GTCAAGATTC	CCGATGCTGC	GACCAACGCC	AAGCTCCGCA	TGCCCATCTC	TGGTGTGATC	[798]
167_EFL_Gracilaria_changii	-----AA	GGTCCACATT	GACACTCTCT	CGACGCCCT	TGAGAAAATG	GTTCAAGATC	CTCTCGTCC	GACTGATAAG	CCCTTGCGA	TGCCCATCTC	TGGTGTGATC	[798]
168_EFL_Cryptomonas_ovata	-----ACTTC	GGTCAAGATC	GTGACTCTCC	TCAACGCCCT	CAACGACTTC	GCTCAGCGCC	CGCTCGTAT	CACCGACGCG	CGCATGCGCT	GCCCGATTTT	TGGATCTCTAC	[756]
169_EFL_Guillardia_theta	-----ACCGAA	GATGAAGATC	GAGACCCCTC	TCCATGCTCT	CAACGACTTT	GCCCGCCCCC	CAAAGCGTAA	CGTGCATGCC	CCCATGCGCT	GCCCATCTC	CGGATCTCTAC	[801]
170_EFL_Rhodomonas_salina	-----GGCAA	GATGACCATC	GTCTGCCCTA	AGGACGCCCT	CAACGACATG	GCCAGCGCGC	CCGAGCGTAT	CACCGACGCT	CCCATGCGCT	GCCCATCTC	CGGATCTCTAC	[798]
171_EFL_Pavlova_lutheri	-----AA	GGTGCACGTC	CACACGCTCT	ACGACGCCCT	CAACGACTTC	GTCAGACGCG	CGAGCGCAA	GACCGACGCG	CGCATGCGCC	TGCCCATCTC	TGGGTGCTAC	[798]
172_EFL_Chrysochromulina_sp.	-----AC	CATGCACATC	GACACCCCTG	ACGACTGCTT	CAACGACATG	TGCAAGCCGC	CCGAGCGTAA	CGACGCCGCC	GCCATGCGCA	TGCCCGTGTG	GGGATCTCTAC	[618]
173_EFL_Isochrysis_galbana	-----GAGAA	GCTGCAGCTC	CACACCCCTG	ACGACTGCTT	CAACGACTTC	TGCACTCCGC	CGGCGCGCAA	CGAAACCGCG	CCCATGCGCA	TGCCCGTGTG	GGGATCTCTAC	[801]
174_EFL_Emiliania_huxleyi	GCAAGACGAA	GATGCACGTC	GACACGCTCT	AGGAGTGCTT	CAACGACTTC	TGCTCGCCCC	CGTCCGCAAC	CGACAAGCGC	GAGATGCGGA	TGCCCGTGTG	GGGATCTCTAC	[807]
175_EFL_Achnanthes_kuwaitensis	-----ACCAA	GATCACTGGA	TTCAACCTTG	TCGAGTGCTT	GAACCACTTC	ATCAGCCCTC	CCAAGCGTTT	CCCCGATCTT	CCGCTCCGTC	TTCCCATCTC	CACCATCTAC	[720]
176_EFL_Asterionella_glacialis	-----GTCAA	GGTTACCGGT	TTCAACCTTC	TCGATGCCCT	CAACAACCTC	ATCAACCTTC	CAAAGCGTGA	ACCCGATGCC	CTCTCTGCTC	TGCCCATTTT	CACCATCTAC	[720]
177_EFL_Detionula_confervacea	-----AAGCA	GGTGACCGGA	TTCAACCTGC	TTGATGCCCT	CAACAACCTC	ATCAACCTTC	CTATCCGCTA	CCCTGAGCTT	CCCGTCCGTC	TTCCCATCTC	CACCATTTTAC	[720]
178_EFL_Thalassiosira_pseudonana	-----ACCGT	TGTCAAGGGA	TTCAACCTTG	TTGATGCCCT	GAACAACCTT	ATCAACCTTC	CCAAGCGTGA	CCCTGAGCTT	CCCGTCCGTC	TTCCCATTTT	CACCATTTTAC	[783]
179_EFL_Perkinsus_marinus	CTAAGGAGTC	GTTCATATCT	GATACCCCTC	TAGACGCTCT	CGAGAAGTTC	GCTACCGTTC	CTAAGCGTGT	GACTGATGCA	CCAATGAGAA	TGCCCGTCTC	TGGTATCTAC	[816]
180_EFL_Oxyrrhis_marina	--TCGGAGAA	GATCCACGTC	CACACCTTGT	ATGATGCTTT	GGACAAGTTC	TGCCCGTGTG	CGGTCGCTCC	CTCTGAGGCG	CGCATGAGGA	TGCCCGTGTG	TGGATCTCTAC	[810]
181_EFL_Karlodinium_micrum	-----AC	CATCCATGTC	GACACAATGT	ATGATGCTTT	GGACAAGATG	TGCAAGATTC	CAGAGCGACC	CTGAGTGCA	CCAATGCGTA	TGCCCATCTC	AGGAATCTAC	[804]
182_EFL_Amphidinium_carterae	-----AC	ATTGCACATC	GACACCCCTG	ATGAAGTGTCT	TGACAAGATG	TGCAAGATTC	CAGAGCGCCC	TGTGAGCGCA	CCAATGCGCA	TGCCCATCTC	TGGATCTCTAC	[804]
183_EFL_Heterocapsa_triquetra	-----AC	CATCCACGTC	GACACGGCTC	ATGATGTGCT	GGACAAGCTC	GTCGCGTCTC	CGGAGCGCCC	CACGAGTGC	CCCATGCGCA	TGCCCATCTC	CGGTATCTAC	[804]
184_EFL_Planoglabratella_operculis	ACAAGATTGT	CGTCAAGGTA	CACACTTTAA	TTGATGCCCT	GGAGAAGTGT	GTGAAGCCAC	CAAAGCGAGC	CAACAAAAAA	GATTTAGAA	TGCCCTGATC	GGGATCTCTAC	[639]
185_EFL_Reticulomyxa_filosa	-----GATTA	GCTTACTGGT	TTGACGTTGA	TGGATGCTTT	GGAGAAGACC	GTCAACGAGA	CAAAAAGACC	CAATTCGTA	TGCCCATCTC	TGCCCATCTC	TGGTGTCTAC	[714]
186_EFL_Lotharella_amoebiformis	-----AAGGA	GTACACCGGT	AAGACCCCTC	TTGAGTGCTT	CGATGAGTTC	GTGAGCGCGC	TGCCCGGTA	GCTTGACAAG	AAGCTTCTG	TGCCCATCTC	TGGCGTTTAC	[870]
187_EFL_Gymnochlorella_stellata	-----GTTAA	GGTCAAGGTA	CAAACTCTCT	TCGAGGCTCT	CGATGAGTTC	GTCGAAACCTG	TAACACGTA	CTTGAGCAAG	CCTCTACGTC	TACCTTTTCT	CGGTGTCTAC	[828]
188_EFL_Chlorarachnion_reptans	-----GTCAA	GGTAAAGGCG	GACACCCCTC	TTGAGGCCCT	TGACGAGTTT	GTGAGACGTC	TGAAGAGTGA	CTTGAGAAAG	CTCTCCGCTC	TTCCGCTCTC	TGGTGTCTAC	[759]
189_EFL_Diplonema_ambulator	-----GAGAA	GCTCCACCTG	GAGACCCCTG	TGACGCCCT	TGACCTGATG	GTGAGGATGT	CGGAGAGGAA	GAAACGATCC	CCCATGAGGA	TGCCCTGTC	TGGCGTGTAC	[402]
190_EFL_Diplonema_papillatum	-----GAGAA	GCTGCACGTC	GACACCCCTC	TGACGCCCT	GAACAACATG	GTGCTCTCTC	CGGAGGGGAA	GAAACGACGC	CCCATGAGGA	TGCCCTGTC	TGGCGTGTAC	[774]
191_EFL_Rhynchopus_euleeides	-----AAGTCGAA	CCTGCACATC	GAGACCCCTG	TGACGCCCT	GAACGACATG	GTGCGCATGT	CTGAGCGCAA	GAAACGACGC	GCCATGCGCA	TGCCCGTGTG	CGGCGTGTAC	[417]
192_EFL_Petalomonas_cantuscygni	-----TCCAA	GCTCCACGCG	GACACCCCTC	TCGACTGCTT	TGACAAGTAC	GTTGAGCTCT	CCCCCGCAA	CGCATGCGCA	TGCCCGTGTG	CGGCGTGTAC	GGGATCTCTAC	[645]
193_EFL_Trypanoplasma_borrelii	GGGGAACCGT	TGTGCACGTC	GACTGCCCTG	TCGACGCACT	GAACAACATG	GTGAACCTGC	CTCCCGTGA	CAATACCGCC	CCCATGCGCG	CGCGTGTGTC	GCAAGTGTAC	[825]
194_EFL_Neobodo_salienis	--AAGAAAGGA	CATCAAGGTC	AACTGCCCTG	TGGACGCCCT	GAACGACATG	GTCAACCTGC	CGCCCGCAA	CACCGAGGCT	GCCCTCCGCG	CCCCGTCGCG	CAAGTGTCTAC	[768]

	1000	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	
001_aEF1a_Desulfurococcus_mobilis	---AATATCT	CTGGCATAGG	CGTTGTGCCA	GTGGGCAAGG	TTGAGACAGG	TGTACTGAAG	GTGGGCGACA	AACATAGTGT	CATGCCAGCC	GGGCTCGTGG	CC-----	[828]
002_aEF1a_Sulfolobus_solfataricus	---TCAATCT	CTGGAGTAGG	TACTGTACCG	GTAGGTAGAG	TAGAAGCGG	AGTGTTAAGG	GTAGGTGATA	AGATAGTATT	CATGCCAGCA	GGCAAGATAG	GT-----	[822]
003_aEF1a_Thermoplasma_acidophilum	---TcCaTaA	cCgGtAtCgG	aAcGgTtCcT	gTcGgCaGaG	tAgAgAcGgG	tGtCcTtAaG	cCcGgTgAcA	aGgTcAtAtT	cCtCcCgGcA	gAcAaGcAgG	gT-----	[804]
004_aEF1a_Halobacterium_sp.	---ACGATCT	CCGGGATCGG	TACCGTCCCG	GTCCGACGCA	TCGAGACCgG	TGTGATGAAC	ACCggGcGAcA	ACGTcAGCTT	CCAGCCcCAGC	GAGCTCGGTG	GC-----	[795]
005_aEF1a_Methanocaldococcus_jannaschii	---TCAATTA	CAGGGTTCGG	AACCTGTCCCA	CTTGGAAAGG	TCGAAACAGG	TATCTTAAAG	CCAGGAGACA	AGATTGTCTT	CGAACCAGCA	GAGATTAGCG	GA-----	[810]
006_EF1a_Acanthamoeba_culbertsoni	---AAGATCG	CGCGTATCGG	TACCGTGCCC	GTCCGCGCTA	TCGAGACCgG	TATCCTCAAG	CCCGGTATGG	TGGTGACCTT	CGCCCCCGCT	AACGTGACCA	CT-----	[792]
007_EF1a_Neoparamoeba_branchipila	---AAGATTG	GTGGTATYGG	CACAGTGCCC	GTCCGTCGTG	TCGAGACTGG	TGCTTTGAAG	CCCGGTATGG	TTGTCACTT	CGCTCCCCC	ATGATCACCA	CT-----	[777]
008_EF1a_Planoprotostelium_aurantium	---AAGATCG	GAGGTATTCGG	AACGGTACCC	GTCCGTCGTG	TTGAGACCgG	TATCTTGAAG	CCCGGAGTCT	CCGTCACTT	CGCCCCCGCC	GGAGTTACCA	CT-----	[774]
009_EF1a_Physarum_polycephalum	---AAGATCG	GAGGTATTCGG	AACGTGCCCC	GTCCGCGCTG	TTGAGACTGG	TATCCTCAAG	CCCGGCATGA	TCGTCACTT	CGCTCCCGCC	AACCTGTCCA	CT-----	[774]
010_EF1a_Amoebidium_parasiticum	---AAGATCG	GTGGCATCGG	CACGTGTGCC	GTGGGCGCTG	TGGAGACCgG	TGTCATCAAG	CCCGGCATGG	TCGTCACTT	CGCCCCCACC	GGTGTGACCA	CT-----	[825]
011_EF1a_Capsaspora_owczaraki	---AAGATTG	CGCGTATTCGG	CACGTGTCCC	GTCCGCGCTG	TCGAGACTGG	TGTCCTCAAG	CCCGGCATGG	TTGTCACTT	CGCCCCCAGC	AACGTGACCA	CC-----	[855]
012_EF1a_Monosiga_ovata	---AAGATCG	GTGGTATTCGG	CACGGTGCCC	GTCCGTCGTG	TTGAGACCgG	TACCTGAAG	CCCGGCATGG	TTGTCACTT	CGCCCCCGGC	AACGTTACCA	CC-----	[834]
013_EF1a_Salpingoeca_amphoridium	---AAGATCG	GTGGCATCGG	CACAGTTCCT	GTCCGTCGTG	TTGAGACTGG	TGTCCTGAAG	CCCGGCATGG	TCGTCACTT	TGCCCCCGGC	AACGTGTGCA	CT-----	[837]
014_EF1a_Ceanorhabditis_elegans	---AAGATCG	GAGGAATCGG	AACGTGTCCA	GTCCGACGTG	TTGAGACCgG	AATCATCAAG	CACGGAATGG	TCGTTACCTT	CGCTCCACAA	AACGTGACCA	CT-----	[861]
015_EF1a_Drosophila_melanogaster	---AAAAATTG	CGCGTATTGG	AACAGTACCC	GTGGGTCGTG	TGGAGACTGG	TGTGCTGAAG	CCCGGTATGG	TTGTGGTCTT	CGCCCCCTGT	AACATCACCA	CT-----	[861]
016_EF1a_Ornithorhynchus_anatinus	---AAGATTG	GAGGCATCGG	GAGGACCGGG	GTGGGCGGGG	TGGAGACCgG	GATCCTGAGA	CTGGGATGG	TAGTGACCTT	TGCCCCGGTG	AACATCACCA	CA-----	[861]
017_EF1a_Xenopus_laevis	---AAAAATTG	GTGGAATTGG	CACGTGTGCC	GTGGGACAGG	TGGAGACTGG	TGTCTTAAAG	CCAGGCATGA	TTGTGACTT	TGCTCCAAGT	AACGTCACTA	CA-----	[861]
018_EF1a_Saccharomyces_cerevisiae	---AAGATCG	GTGGTATTGG	TACTGTGCCA	GTCCGTAGAG	TTGAAACCgG	TGTCATCAAG	CCAGGTATGG	TTGTTACTT	CGCCCCAGCT	GGTGTTACCA	CT-----	[855]
019_EF1a_Neurospora_crassa	---AAGATCG	GTGGTATTCGG	CACAGTGCCC	GTCCGCGCTA	TCGAGACTGG	TGTCCTCAAG	CCCGGTATGG	TCGTTACCTT	CGCTCCTTCC	AACGTGACCA	CT-----	[858]
020_EF1a_Cryptococcus_neoformans	---AAGATCG	GTGGTATTCGG	CACAGTCCCT	GTCCGCGGAG	TCGAGACCgG	TGTCATCAAG	CCCGGTATGG	TCGTCAAGTT	CGCCCCCACC	AACGTGACCA	CT-----	[855]
021_EF1a_Schizophyllum_comune	---AAGATCG	GTGGTATTGG	GACGGTGCCC	GTCCGCGCTG	TCGAGACCgG	TATCATCAAG	CCCGGCATGG	TCGTCACTT	CGCCCCCTCC	AACGTGACCA	CT-----	[855]
022_EF1a_Monoblepharis_insignis	---AAGATCG	GTGGTATTCGG	TACTGTCCCT	GTGGGTCGTG	TCGAGACTGG	TGTGATCAAG	CCCGGTATGG	TCGTTCACCT	CGCTCCCTCG	AACGTGACCA	CT-----	[741]
023_EF1a_Rhizophlyctis_harderi	---AAGATTG	GTGGTATTGG	AACAGTTCCT	GTCCGTCGTG	TTGAAACTGG	TACCATCAAG	CCGTGATGA	TTGTCAACTT	TTCCCTTGCC	AACGTTACCA	CT-----	[741]
024_EF1a_Harpochytrium_sp.	---AAGATTG	CGCGTATTCGG	AACAGTACCC	GTCCGTCGTG	TTGAGACTGG	TATCATCAAG	CCCGGTATGG	TCGTGACCTT	CGCCCCCACC	GGTGTGACCA	CT-----	[693]
025_EF1a_Absidia_glauca	---AAGATCG	GTGGTATTGG	TACTGTGCC	GTCCGTCGTG	TCGAAACTGG	TGTCATCAAG	CCCGGTATGG	TCGTCACTT	CGCTCCCGCT	AACGTGACCA	CT-----	[855]
026_EF1a_Smittium_simulii	---AAGATTG	GTGGTATTCGG	AACAGTTCCT	GTCCGTCGTG	TTGAAACTGG	TGTCATTAAG	CCCGGTATGA	TCGTCAACTT	TGCCCCCATCA	TATGTTACTA	CT-----	[828]
027_EF1a_Chara_australis	---AAGATCG	CGCGTATTGG	CACGGTGCCC	GTGGGAAAGG	TTGAGACTGG	TATAATTAAAG	CTTGGCATGG	TGGTGACCTT	CGCGCCTTCT	GGATTGACGA	CT-----	[774]
028_EF1a_Physcomitrella_patens	---AAGATTG	GAGGCATCGG	AACCTGTGCCA	GTGGGACGGG	TGGAGACAGG	AATCATCAGG	CCAGGTATGT	TGGTGTGCTT	CGACCCGACG	GGACTGACGA	CT-----	[825]
029_EF1a_Acetabularia_acetabulum	---AAAAATTG	CGCGAATTGG	AACCGTGCCC	GTCCGGAAGG	TGGAGACAGG	TGTGATCAAG	CCCGGAATGA	TTGTGACCTT	TGCCCCCACA	GGTTTGACGA	CG-----	[774]
030_EF1a_Chaetomorpha_coliformis	---AAGATCG	GTGGAATCGG	AACAGTGCCC	GTGGGCGCTG	TGGAGACTGG	TGTCATTAAG	CCCGGTATGG	TTGTGGACTT	CGCCCCCTGT	GGACTCACCA	CT-----	[624]
031_EF1a_Arabidopsis_thaliana	---AAGATTG	GTGGTATTGG	AACGGTGCCA	GTGGGACGTG	TTGAGACTGG	TATGATCAAG	CCGTGATGG	TTGTTACCTT	TGCTCCACCA	GGGTGACCA	CT-----	[825]
032_EF1a_Oryza_sativa	---AAGATCG	CGCGTATTGG	TACCGTGCTT	GTGGGTCGTG	TTGAGACTGG	TGTCTCAAGT	TCGTGATGG	TGGTGACCTT	TGGTCCAGC	GGTCTGACCA	CT-----	[825]
033_EF1a_Cyanidioschyzon_merolae	---AAGATCG	CGCGTATTGG	CACGGTGCCC	GTCCGTCGGG	TAGAGACGGG	TATCCTGCGG	CCAGGCATGG	TGGTTACCTT	CGCGCCCACT	GGTCTGACGA	CT-----	[837]
034_EF1a_Porphyrha_yezoensis	---AAGATCG	CGCGCATGGG	CACAGTGCCC	GTGGGCGCTG	TGGAGACTGG	GTCATCATCAAG	CCCGGCATGG	TCGTGACCTT	TGCGCCCTCG	GGCTCTGCGA	CT-----	[837]
035_EF1a_Plasmodium_falciparum	---AAAAATCG	TTGGTATTGG	TACCGTGCTT	GTGGTCTGCT	TTGAAACAGG	TATCTTAAAG	CCCGGTATGG	TTTAAACTT	CGCTCCATCA	GGCTTTGTGT	CT-----	[819]
036_EF1a_Cryptosporidium_parvum	---AAGATAG	GTGGTGTAGG	TACTGTCCCA	GTCCGTCGTG	TTGAGACTGG	TATCATCAAG	CCAGGTATGA	ATGTTACCTT	CGCTCCAGCT	GGTGTTACCA	CT-----	[813]
037_EF1a_Tetrahymena_pyriformis	---AAGATCG	GTGGTATTCGG	AACAGTGCCC	GTCCGTAGAG	TCGAAACTGG	TGTCATCAAG	CCCGGTATGT	CCATTCAATT	TGCCCCTAAC	GAAGTATGCT	CT-----	[828]
038_EF1a_Stentor_coeruleus	---AAGATTT	CTGGTATTGG	CACAGTCCCA	GTCCGCGCTG	TTGAAACTGG	TTCAATTAAAG	CGTGGTATGG	TCGTTGTTTT	TGCCCCCATCA	GGAGTTACTA	CT-----	[780]
039_EF1a_Phytophthora_parasitica	---AAGATCG	CGCGTATTGG	CACGGTACCT	GTCCGCGCTG	TGGAGACCgG	TGTCATCAAG	CTTGGCATGG	TGCGCACTT	TGCCCCCGTT	GGTCTGTGCA	CT-----	[825]
040_EF1a_Mallomonas_rasilis	---AAAAATTG	GAGGTATTCGG	AACGTGCTCT	GTCCGACGTG	TTGAGACTGG	AATCCTCAAG	CCTAACATGG	TTGTCACTT	CGCCCCGTT	GGGCTCACCA	CT-----	[786]
041_EF1a_Phaeodactylum_tricornutum	---AAGATCG	GTGGTATTCGG	AACAGTCCCG	GTCCGCGCGG	TCGAGACCgG	AGTCATCAAG	CTCGGTATTC	ACGCTATGTT	CGCCCCGCTC	GGTATTATTG	CC-----	[825]
042_EF1a_Thalassiosira_pseudonana	---AAGATTG	GAGGCGTTCG	AACAGTTCCT	CGCGGAAAGG	TTGAGACCgG	TGTGATGAAG	CCGGGCATGC	ACATAGAGTT	TGCTCCAAT	GGCATCTGGG	CT-----	[825]
043_EF1a_Malawimonas_jakobiformis	---AAGATTG	CGCGTATTCGG	CACAGTGCCC	GTGGGTCGTG	TCGAGACCgG	TATCTTGAAG	CCCGGTATGG	TGGTGCACTT	CGCCCCGCTC	GGGCTGACCA	CT-----	[756]
044_EF1a_Malawimonas_californiana	---AAGATCG	GTGGCATCGG	CACGTGTGCC	GTGGGCGCTG	TCGAGACCgG	TATCATCAAG	CCCGGCATGG	TGGTCACTT	TGCCCCCGCT	GGGCTGACCA	CT-----	[825]
045_EF1a_Trimastix_pyriformis	---AAGATTG	CGCGTATTCGG	GACTGTGCCC	GTCCGCGCTG	TTGAGACCgG	TATCCTCCG	CCCGGCGAGG	TTATCACCAT	TGCCCCCGCC	AACCTGTCCA	CT-----	[783]
046_EF1a_Dinenympha_exilis	---AAGATTG	GTGGTATTGG	AACAGTGCTT	GTGGTCTGCT	TGGAGACTGG	TATCCTACC	CTGGGCGAGG	TTATCACGAT	TGCACCTGTC	ATGATAACAA	CG-----	[765]
047_EF1a_Monocercomonoides_sp.	---AAGATCG	GAGGTATTCGG	AACAGTCCCT	GTCCGTCGTG	TTGAAACTGG	TGTTTTGACT	CCAGGTATGG	TTGTCACAAT	TGCACACGCT	GCAATTACAA	CT-----	[825]
048_EF1a_Streblospio_strix	---AAGATTG	GTGGAATTGG	AACAGTGCCA	GTGGTCTGCT	TGGAAACAGG	AGTCTTGACT	CCAGGACAAAG	TTGTAACAT	TGCACACGCT	ATGATCACAA	CA-----	[774]
049_EF1a_Histomonas_meleagridis	---AAGATCT	CTGGTATTCGG	TACAGTTCCT	GTGGTCTGCT	TTGAATCCGG	TATCATGAAG	CCAAACATGA	CAGTCGCTCT	CGACCATCA	ACAGTACCA	CC-----	[825]
050_EF1a_Trichomonas_vaginalis	---AAGATCA	ACGGTATTCGG	TACAGTTCCT	GTCCGCGCTG	TCGAATCCGG	CACAATGAAG	CACGGATGA	TCGTTAACTT	CGCCCCATCC	ACAGTTACAG	CT-----	[825]
051_EF1a_Hexamita_inflata	---AAGATCA	ACGGTGTTCG	TACCGTCCCA	GCTGGTCTG	TCGAATTCGG	TGAACCTATC	CCAGGTATGA	TGGTGTCTT	CGCCCCAGCC	GGCGAAAAGA	CT-----	[777]
052_EF1a_Spiroplasma_vortens	---AAGATCA	ACGGTGTTCG	TACCGTCCCG	GCTGGGCGCG	TCGAATTCGG	CTGGCTGAAG	CCCAACACCC	TGCGGCTCTT	CGCCCCGACC	AACACCCAGC	CT-----	[777]
053_EF1a_Andalucia_incarcerata	---AAGATCG	GTGGTATTGG	AACAGTGCCC	GTGGTCTGCT	TCGAAACTGG	CAAGCTGGAG	GCTGGTATGA	CTGTTGCTT	CGCTCCCTCT	GGTGTGACCA	CT-----	[774]
054_EF1a_Seculamonas_ecuadoriensis	---AAGATTG	CGCGTATTCGG	CACGTGTGCT	GTGGGCGCTG	TCGAGACCgG	TATCATGAAG	CCGGGCATGA	CTGTGCTGTT	CGCGCGCTG	GGTCTCTGCA	CC-----	[825]
055_EF1a_Reclinomonas_americana	---AAGATTG	CGCGTATTCGG	CACGGTGCCC	GTCCGCGCTG	TGGAGACTGG	TATCCTGCG	CCGGGCATGG	TCGTGACCTT	CGCCCCCTGT	GGGCTGACCA	CT-----	[825]
056_EF1a_Histiona_roides	---AAGATTG	CGGGTATTCGG	CACGTGTGCCA	GTGGGCGCTG	TCGAGACTGG	TGTGATGAAG	CCAGGCATGA	CTGTGCTCT	CGCCCCATGC	GGTCTCTGCA	CT-----	[825]
057_EF1a_Jakoba_libera	---AAGATCG	GTGGTATTCGG	CACGTGTGCC	GTGGGCGCTG	TCGAGACCgG	TATCATGAAG	CCGGGCATGG	TCATCACTT	CGCCCCCTCC	GGTGTGACCA	CT-----	[825]
058_EF1a_Stachyamoeba_lipophora	---AAGATCT	CTGGTATTCGG	CACGTGTACC	GTCCGTCGTG	TTGAGACTGG	TATCCTCAAG	CCCGGTATGA	ACGTTACCTT	CGCCCCCTGT	GGTGTCTCCA	CT-----	[822]
059_EF1a_Naegleria_gruberi	---AAGATTG	GTGGTATTGG	TACTGTCCCA	GTCCGTAGAG	TTGAAACTGG	TAAATTTGAAG	CACGATATGA	TGGTTCACTT	TGCTCCAGGT	CAAGCTGTCTG	ATACT----	[777]
060_EF1a_Naegleria_andersoni	---AAGATTG	GTGGTATTGG	TACTGTCCCA	GTCCGTAGAG	TTGAAACTGG	TAAATTTGAAG	CCAGGTATGA	TGATTCACCTT	TGCTCCAGGT	CAAGCTGTCTG	ATACT----	[795]
061_EF1a_Acrasis_rosea	---AAAAATTG	GTGGTATTCGG	TACCGTCCCA	GTCCGTCGTG	TTGAGACTGG	TGTAATGAAG	CACGGTATGA	CCGTTCTCTT	CGCACAGTGT	GAGCTCTCAA	CT-----	[774]
062_EF1a_Sawyeria_marylandensis	---AAAAATTG	GTGGTATTGG	TACTGTTCCT	GTAGGTAGAG	TTGAAACTGG	TGTCATGACT	CCAGGTATGA	CGGTTTACTT	CTCACAGCA	GCAAAATCAA	CT-----	[825]
063_EF1a_Bodo_saltans	---AAGATCG	CGGGTATTCGG	CACCGTTCCT	GTCCGCGCTG	TTGAGACTGG	TATCCTCCG	CCCGGTATG	TCGTCACTT	CGCCCCCGCC	AACATCACCA	CT-----	[765]
064_EF1a_Rhynchobodo_sp.	---AAGATCG	GTGGTATTCGG	TACCGTCCCA	GTCCGCGCTG	TCGAGACCgG	CATCATGAAG	CCCGGTATG	TCGTGCTCT	CGCCCCCGCT	GGGCTGACCA	CT-----	[759]
065_EF1a_Rhynchomonas_nasuta	---AAGATCG	CGGGCATTCGG	CACCGTGCCC	GTCCGCGCGG	TCGAGACCgG	CATCATGAAG	CCCGGCGACG	TCGTCACTT	CGCCCCCGCC	AACGTTCTCTA	CT-----	[822]
066_EF1a_Dimastigella_trypaniformis	---AAGATCG	CGGGCATTCGG	CACCGTGCCC	GTCCGCGCTG	TTGAGACCgG	CATCCTGAAG	CCGGGTGACG	TCGTCACTT	CGCCCCCGCC	GGGCTGACCA	CC-----	[780]

	1000	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	
133_EFL1a_Trachelomonas_volvocina	---AAAATTG	GAGGTATTGG	AACCTGTGCTT	GTTGGTCGTG	TGGAACACAGG	TGTTCTAAAG	CCTGGTGAAG	TGGTCACCTT	CGCCCAAAAT	GCCTTGACAA	CT-----	[822]
134_EFLenvironmental_sample_EU810232	---GCCATCA	AGGGAACTGG	TGACGTGCTG	ACCGGACGCG	TTGAGCAGGG	ATCGCCAAAG	CCCGGTGACG	AGGTGGTCTT	CCTGCCCAAC	CACACGCCCG	CCAAACCCCTG	[905]
135_EFLCreolimax_fragrantissima	---AAGATCA	AGGGTGTGGG	AGATGTTCTT	ACCGGACGCTG	TCGAGCAGGG	AACCGTCAAG	CCTGGAGATG	AGTGATCTTT	CTTGCCCTACC	CACACTCCCT	CCACCGGCTG	[833]
136_EFLSphaeroforma_arctica	---AAGATCA	AGGGTGTGGG	AGATGTTCTA	ACTGGACGCTG	TTGAGCAGGG	AACCGTAAAG	CCCGGAGATG	AGTTGTTTTT	CCTACCCACC	CATACCCCCCT	CCACTGGTTG	[905]
137_EFLMonosiga_brevicollis	---AAGATCA	AGGGTGTGGG	AGATGTGCTC	ACTGGCCGCTG	TCGAGCAGGG	CACCGTCAAG	CCCGGAGATG	AAAGTTGCTT	CCTGCCCAAC	CACACTTCTC	CCACCGGCTG	[905]
138_EFLXenopus_laevis	---AAGATCA	AGGGTGTGGG	CGATGTGCTG	ACGGGCGCTG	TCGAGCAGGG	CACCGTCAAG	CCCGGTGATG	AGGTGATCTT	CCTGCCCAAC	CACACCCCTC	CCACCGGCTG	[905]
139_EFLSpizellomyces_punctatus	---AAGATCA	AGGGTGTGGG	TGATGTGCTC	GCTGGCCGCTG	TTGAGCAGGG	TACCGTCTTG	CCCAACAAGG	ATGTTGCTTT	CGTCCCCACC	CACACCGCTG	CCAAACCATG	[911]
140_EFLGaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
141_EFLAllomyces_macrogynus	---AAGATCA	AGGGTGTGGG	TGATGTACTT	GCCGGCCGCTG	TCGAGCAGGG	TATCGTCAAG	ACCAACAAGG	ATGTCGTCTT	CCTGCCCAAC	CACACGACGG	CCACGGCGTG	[911]
142_EFLBlastocladiella_emersonii	---AAGATCA	AGGGTGTGGG	AGATGTCTCT	GCCGGCCGCG	TCGAGCAGGG	TATCGTCAAC	CCCAACAAGG	ATGTCGTCTT	CCTGCCCAAC	CACACCAACG	CCACCCCTG	[911]
143_EFLCatenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
144_EFLCoelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
145_EFLEntomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
146_EFLBasidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
147_EFLConidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
148_EFLTetraselmis_tetrathele	---AAGATCA	AGGGTGTGGG	TGATGTGCTC	GCCGGACGCG	TGGAGCAGGG	CCAGGTGAAG	CCCAACGATG	AGGTTGCTTT	CATGCCCAAC	CACACCAAGG	CCAACCTCTG	[854]
149_EFLMesostigma_viride	---AAGATCA	AGGGTGTGGG	TGATGTGCTG	GCCGGCCGCTG	TCGAGCAGGG	AATCGTCAAG	CCTGGTGAAG	AAGTCTCTG	-----	-----	-----	[879]
150_EFLOstreococcus_tauri	---AAGATCA	AGGGTGTGGG	TGATGTCTCT	GCCGGCCGCG	TCGAGCAGGG	CCTCGTCAAG	CCGGGTGAAG	ATGTCGTCTT	CCTGCCGACC	CACACCTCCT	CCAACCCGCTG	[914]
151_EFLNephroselmis_olivacea	---AAGATCA	AGGGTGTGGG	TGATGTGCTG	GCCGGACGCG	TGGAGCAGGG	CACCGTCAAG	CCCAACAAGG	AGGTGATCTT	CCTGCCAACC	CACACCGCCT	CCAAGCCCTG	[881]
152_EFLMicromonas_pusilla	---AAGATCA	AGGGTGTGGG	TGATGTCTCT	GCCGGCCGCG	TYGAGCAGGG	CGTCTGCCG	CCCGGTGAGG	AGGTGATCTT	CATCCCAACC	CACACCGCTC	CCAAGCCCTG	[863]
153_EFLParachlorella_kesslerii	---AAGATCA	AGGGTGTGGG	TGATGTGCTG	GCCGGCCGCG	TGGAGCAGGG	CATTGTCAAG	CCCGGTGAGG	AGGTTATCTT	CCTGCCCAAC	CACACCGCCT	CCAAGCCCTG	[866]
154_EFLHelicosporidium_sp.	---AAGATCA	AGGGTGTGGG	TGATGTGCTC	GCTGGCCGCG	TGGAGCAGGG	CATCATCAAG	CCCAACGAGG	AGGTTGCTTT	CATGCCCAAC	CACACCGCCT	CCAACCAAGT	[926]
155_EFLGonium_pectorale	---AAGATCA	AGGGTGTGGG	TGATGTGCTG	GCTGGCCGCG	TGGAGCAGGG	CGTCTCAAG	CCTGGTGAAG	AGGTGGTCTT	CCTGCCCAAC	CACACCAACG	CTAACCCCTG	[908]
156_EFLChlamydomonas_reinhardtii	---AAGATCA	AGGGTGTGGG	TGATGTGCTG	GCCGGCCGCG	TCGAGCAGGG	CGTGGTCAAG	CCCGGTGATG	AGGTGATCTT	CCTGCCCAAC	CACACCAACG	CCAAGCCCTG	[908]
157_EFLScenedesmus_obliquus	---AAGATCA	AGGGAGTGGG	TGATGTGCTG	GCCGGCCGCTG	TGGAGCAGGG	CGTGGTCAAG	CCCGGTGATG	AGGTGATCTT	CATGCCCAAC	CACACCAACG	CCAAGCCCTG	[908]
158_EFL"Oryza_sativa"	---AAGATCA	AGGGTGTGGG	TGATGTGCTG	GCCGGCCGCTG	TGGAGCAGGG	TGTTGTCAAG	CCCGGTGATG	AGGTGATCTT	CCTGCCCAAC	CACACCGCCT	CCAAGCCCTG	[908]
159_EFLChlorococcum_sp.	---AAGATTA	AGGGTGTGGG	AGATGTGCTT	GCCGGCCGCTG	TTGAGCAGGG	TATGTTCAAG	CCCGGTGAGG	AGGTTCTCTT	CATGCCCAAC	CACACTGCTG	CCAACAAGTG	[857]
160_EFLOchlochaete_hystrix	---AAGATCA	AGGGTGTGGG	TGATGTGCTG	GCCGGCCGCG	TGGAGCAGGG	TGTTGTCAAG	CCCGGTGATG	AGGTGGTCTT	CATGCCCAAC	CACACCCCTG	AGAACCCCTG	[866]
161_EFLUlva_intestinalis	---AAGATCA	AGGGTGTGGG	AGATGTGCTT	GCCGGCCGCG	TCGAGCAGGG	CATTGTCAAG	CCGAATGACG	AGGTTGCTTT	CATGCCCAAC	CACACCGCTG	CCAACAAGTG	[857]
162_EFLUrospora_sp.	---AAGATCA	AGGGCGTGGG	CGACGTGCTC	GCCGGCCGCG	TCGAGCAGGG	CGTYGTCAAG	GCCAACGACG	AGGTGCGCTT	CATGCCCAAC	CACACCGCCT	CCAACAAGTG	[710]
163_EFLGymnochlorella_stellata_ppc	---AAGATTA	AGGGGATAGG	TGATGTCTCT	GCAAGGCGCTG	TGAGCAGGG	AGTGTAAAG	AATGAAGCTG	ATGATGATCT	CTTCCCAACG	CACGTGCCCTA	ATCATCCCTG	[944]
164_EFLLotharella_amoebiformis_ppc	---CGTATCA	AGGGAGTGGG	GACGTGCTCT	GCAAGGACAG	TAGAGCAGGG	CGTCTCAAG	AATGAAGCCG	ACGTGATATT	CTTCCCAACG	CACGTGCCCTA	ATCATCCCTG	[689]
165_EFLChlorarachnion_reptans_ppc	---CGCATCA	AGGGTGTAGG	CGATGTCTCT	GCAAGGCGCG	TCGAGCAGGG	AGTGTGTCAG	AATGAAGCCG	ATGTCGTCTT	CCTCCCATCA	CACACACCAA	ACAACCCCTG	[716]
166_EFLChondrus_crispus	---AAGATCA	AGGGTGTGGG	TGATGTGATC	ACCGGCGCTG	TCGARYWSGG	YRYCAKCAAG	CCTGAGGAGC	AGGTTCTCTT	CCTGCCGACT	CACACTACCT	CCAACGCCCTG	[905]
167_EFLGracilaria_changii	---AAAATTA	AGGGTGTGGG	TGATGTCTCT	ACTTGTCTTG	TGGAACAAGG	TGCTCTCAAG	CCTGGAGACC	AGATTCTTTT	CTTCCGACT	CACACTTCTC	CCACTCCGCTG	[905]
168_EFLCryptomonas_ovata	---AAAATTA	AGGGCGTAGG	CGACGTGTTG	GCTGGCCGCTG	TCGAGCAGGG	TCTGGTGAAC	CCCGGTGAGG	ATGTCGTTTT	CATGCCCAAC	CACGCTTCGG	GCACCTCCGCTG	[863]
169_EFLGuillardia_theta	---AAGATCA	AGGGTGTGGG	TGACGTGTTT	GCTGGACGCTG	TTGAGCAGGG	AATCGTCAAC	CCCGGAAAGG	ATGTCATCTT	CATGCCCAAC	CACACTCCTG	GTACCCCTTG	[908]
170_EFLRhodomonas_salina	---AAGATCA	AGGGTGTGGG	TGACGTGTTG	GCCGGACGCG	TCGAGCAGGG	TCTCGTGAAC	CCGGGAGAGG	ATGTCATCTT	CATGCCCAAC	CACACCAACG	GTACCCCTG	[905]
171_EFLPavlova_lutheri	---AAGATCA	AGGGCGTGGG	CGACGTCTCT	GCCGGCCGCG	TCGAGCAGGG	CATCGTCAAG	CCCGGCGAGG	ACGTGATCTT	CCTCCCGACG	CACACGACGG	TCAAGCCATG	[905]
172_EFLChrysochromulina_sp.	---AAGATCA	AGGGTGTGGG	TGACGTGATT	GCCGGCCGCA	TCGAGCAGGG	CCTCGTCAAG	CCCAACGAGG	AGGTGATCTA	CATGCCCAAC	CACACCGCGG	CCAACGCCTG	[725]
173_EFLIsochrysis_galbana	---AAGATCA	AGGGTGTGGG	CGATGTGCTG	GCCGGCCGCG	TCGAGCAGGG	CCTCGTGAAG	CCCAACGAGG	AGGTGATCTT	CATGCCCAAC	CACACCGCGG	CCAAGGCCTG	[908]
174_EFLEmiliana_huxleyi	---AAGATCA	AGGGCGTGGG	CGACGTGCTC	GCCGGCCGCG	TCGAGCAGGG	CGTGATCACG	ACCAACACCG	AGGTGATCTT	CATGCCCAAC	CACACCGCAG	CCAAGCCCTG	[914]
175_EFLAchnanthes_kuwaitensis	---AATATTA	AGGGTGTGGG	TCAGATCATC	TGCGGACCGG	TCGAGCAGGG	CACCATCCCG	CCCGGAGATG	TCATCGGAT	GGTCCCCCTC	AACCTTAAGG	GA-----	[819]
176_EFLAsterionella_glacialis	---AATATTA	AGGGTGTGGG	ACAAATATAT	TGCGGTACCG	TCGAACAAGG	AACCTTCTCG	CCCGGTGATG	TCATCGGAAT	TGCCCCCGCT	GGACTCAAGA	AC-----	[819]
177_EFLDetonula_confervacea	---AACATTA	AGGGAGTTGG	ACAGATCATC	TGCGGAACCG	TCGAGCAGGG	AACCTCAGA	CCTGGTGAAG	TCGTGGAAT	TGTTCCCTCT	GGCTCAAGG	GC-----	[819]
178_EFLThalassiosira_pseudonana	---AACATTA	AGGGAGTTGG	ACAGATCATC	TGCGGAACCG	TCGAGCAGGG	AACCTCACC	CCCGGTGATG	TTATCGGAAT	TGTTCCCTCT	AACCTCAAGG	AC-----	[882]
179_EFLPerkinsus_marinus	---AAGATCA	AAGGTGTGGG	AGATGTGATT	ACCGGCGCTG	TTGAGCAGGG	CATCGTCAAG	CCCAACGAGG	AGGTGCTGTT	CATGCCCTACT	CACACGCGCT	CTACTGCTTG	[923]
180_EFLOxyrrhis_marina	---AAGATTA	AGGGTGTGGG	CGATGTGCTG	GCTGGCCGAG	TGGAGCAGGG	AATGTTGAAG	CCGGGTGAGG	AAAGTGTGTT	TCTGCCCAAC	CACACCCCTG	CCAAGCCTTG	[917]
181_EFLKarlodinium_micrum	---AAAATCA	AAGCGCTGGG	CGATGTATTG	GCAAGTCCGCG	TTGAGCAGGG	TATGTTCAAA	CAAGGCGAAG	AAAGTGGTTT	CTTGCCCAAC	CACACAACAT	CGAATCCTTG	[911]
182_EFLAmphidinium_carterae	---AAGATCA	AGGGTGTGGG	TGATGTCTTG	GCTGGTCCGCG	TGGAGCAGGG	CATTGTGAAG	CAAGGTGAAG	AGGTGGTCTT	CTTGCCAACC	CACACAGCAT	CAAACCCCTG	[911]
183_EFLHeterocapsa_papillata	---AAGATCA	AGGGTGTGGG	CGATGTGCTG	GCCGGCCGCG	TCGAGCAGGG	TGTTGTGAAG	CCCGGTGAGG	AGGTGATCTT	CCTGCCGACG	CACACGGCTT	CCAAGCCGCTG	[911]
184_EFLPlanoglabratella_operculis	---AAAATTA	AAGGTGTGGG	TGATGTGATC	ACAGGCGCTA	TCGAACAAGG	TACCAATTAC	CCTGGAGTCA	ATGTGAGATT	CTACCCGACA	GGATGTGGA-	-----	[735]
185_EFLReticulomyxa_filosa	---AACATTA	AAGGAGTGGG	TGATGTGATT	TGGAACAAGG	ACAATAATCC	CCTGGTCTG	CTGTTGCTG	-----	TTATCCAACA	CACGACAAC-	-----	[810]
186_EFLLotharella_amoebiformis	---AAGATGA	AG--GCCGG	TACCATGCTC	ACCGGACGTA	TTGAGCAGGG	TGTTCTTGA	AAGGAGGTCA	AGACCAAGAC	TGGTGTGCT	GGTAACCCCA	TCCGCTTCTA	[974]
187_EFLGymnochlorella_stellata	---AAGATGT	CT--GCTGG	TACCGTCTAC	ACTGGACGTA	TCGAACAAGG	TCTCTTGA	AAGGAAGTGA	AGACCAAGAC	TGGTACTACC	GGTACTCCAG	TGAAGTTCTA	[932]
188_EFLChlorarachnion_reptans	---AAGATGT	CC--GCTGG	TGTTGTAAAT	ACCGGTCGTA	TCGAGCAGGG	TAACTTGAAG	AAGGAAGATTA	AATCTAAGGT	TGGTACCATT	TGGTACCCTG	TGAAGTTTTT	[863]
189_EFLDiplonema_ambulator	---AAGATCA	AGGGTGTGGG	CGATGTGATC	ACCGGTCGCT	TGGAGCAGGG	TGTTGTGAAG	CCTGGTGAAG	AGGTGCTCTT	CCTGCCCAAC	CACACGGCGG	CGAAGCCGCTG	[509]
190_EFLDiplonema_papillatum	---AAGATCA	AGGGTGTGGG	CGATGTGATC	ACTGGCCGCTG	TGGAGCAGGG	AGTTGTGAAG	CCCGGTGAGG	AGGTTGCTTT	CCTGCCGACG	CACACGGCTT	CCAAGCCGCTG	[881]
191_EFLRhynchopus_euleeides	---AAGATCA	AGGGTGTGGG	TGATGTGATC	ACCGGCGCTG	TGGAGCAGGG	TGTTGTGAAG	CCCGGTGAGG	AGGTGCTGTT	CCTGCCCAAC	CACACGGCGG	CCAAGCCGCTG	[524]
192_EFLPetalomonas_cantuscygni	---AAGATCA	AGGGTGTGGG	CGATGTGCTG	ACCGGCGCTG	TCGAGCAGGG	CATTGTCAAC	CCGGGCGATG	AGGTGCTGTT	CCTGCCCAAC	CACACTCCTC	CCAACAAGTG	[752]
193_EFLTrypanoplasma_borrelii	---AGCATCA	AGGGAGTGGG	TGACGTGCTG	ACCATGCGCA	TCGAACAAGG	TCAGTGGTGG	AAGGGGAGTG	AGGTGATCTT	CCTGCCCAAC	CACACTCCTC	CCAACCCATG	[932]
194_EFLNeobodo_salienis	ACCTCGATCA	AGGGATATCGA	GTGCGTGCCT	ACCTCCCGCC	TCGAGCAGGG	TACCCTGCTG	AAGGGCGACG	AGGTGCTCTT	CCTGCCCAAC	CACACCGCCT	CGAAGCCGCTG	[878]

	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210	
001_aEF1a_Desulfurococcus_mobilis	-----G	AGGTCAAGAC	CATAGAGACA	CACCACACTA	AGATAGAGAA	GGCTGAGCCA	GGCGACAACA	TCGGGTTCAA	CGTGAAGGCG	GTGGAGAAG-	[918]	
002_aEF1a_Sulfolobus_solfataricus	-----G	AAGTTAAGTC	AATAGAAACT	CACCACACTA	AGATGGATTA	AGCGGAGCCA	GGTGACAAGT	TTGGATTAA	CGTTAGAGGT	GTGGAAGAG-	[912]	
003_aEF1a_Thermoplasma_acidophilum	-----G	aTgTcAaAtC	CatCgAaAtG	CatCgAaAtG	CActTgCaGcA	gGcAgAgCcT	gGcGaTaAcA	TAAGcTtCaA	CgtCaGaGgC	aTtGcAaAg-	[894]	
004_aEF1a_Halobacterium_sp.	-----G	AGGTCAAGAC	CATCGAGATG	CACCACGAGG	AAAGTCCGAA	C6CGGAAGCC	GGTGACAACG	TCGGGTTCAA	CGTACCGGCG	ATCGGCAAG-	[885]	
005_aEF1a_Methanocaldococcus_jannaschii	-----G	AAGTTAAAGTC	AATTGAAGTC	CACCACGAAC	AAATTCCACA	AGCAGAACCA	GGAGACAACG	TTGGATTAA	CGTTAGAGGA	GTGGAATAAG-	[900]	
006_EF1a_Acanthamoeba_culbertsoni	-----G	AAGTCAAGTC	GGTCGAGATG	CACCACGAGG	CCATGCCTGA	GGCCGTGCC	GGTGACAACG	TCGGCTTCAA	CGTTAAGAAC	GTCTCCATC-	[882]	
007_EF1a_Neoparamoeba_branchipila	-----G	AGGTCAAGTC	CGTCGAGATG	CATCAGAGC	AAATGAAGCA	AGCCGTCCCT	GGTGACAACG	TCGGATTCAA	CGTCAAGAAC	GTCTCCGTC-	[867]	
008_EF1a_Planoprotostelium_aurantium	-----G	AGGTCAAGTC	CGTCGAGATG	CACCACGTCC	AGCTCCCCGA	GGCCAGACCC	GGAGACAACG	TTGGTTTCAA	CGTCAAGAAC	GTTCCTGTC-	[864]	
009_EF1a_Physarum_polycephalum	-----G	AAGTCAAGTC	CGTCGAGATG	CACCACGTCC	CCCTCCCTGA	GGCCGTCCCA	GGAGACAACG	TCGGTTTCAA	CGTGAAGAAC	TTGTCTGTCT-	[864]	
010_EF1a_Amoebidium_parasiticum	-----G	AAGTTAAGTC	CGTGAGATG	CATCAGGAAG	CTCTCTCCGA	A-----					[867]	
011_EF1a_Capsaspora_owczaraki	-----G	AAGTCAAGTC	GGTCGAAATG	CACCACGAGT	CCCTCCCCGA	GGCCAAACCC	GGTGACAACG	TCGGCTTCAA	CGTCAAGAAC	GTTCGCCGTC-	[945]	
012_EF1a_Monosiga_ovata	-----G	AGGTCAAGTC	TGTTGAGATG	CACCACGAGC	AGCTGCAAGG	GGCCGTCCCC	GGTGACAACG	TCGGCTTCAA	CATCAAGAAC	GTCTCCGTC-	[924]	
013_EF1a_Salpingoeca_amphoridium	-----G	AAGTCAAGTC	TGTTGAGATG	CACCACGAGC	AGCTCGTCTGA	GGCCGTGCC	GGCGACAACG	TTGGCTTCAA	CGTCAAGAAC	TTGTCCGTC-	[927]	
014_EF1a_Ceanorhabditis_elegans	-----G	AAGTCAAGTC	CGTTGAGATG	CATCAGGAAT	CTCTTCCAGA	GGCCGTCCCA	GGAGACAACG	TTGGATTCAA	CGTCAAGAAC	GTCTCCGTC-	[951]	
015_EF1a_Drosophila_melanogaster	-----G	AGGTCAAGTC	CGTGAGATG	CACCACGAGG	CCCTGCAGGA	GGCCGTTCCT	GGAGACAACG	TTGGCTTCAA	CGTCAAGAAC	GTGTCCGTC-	[951]	
016_EF1a_Ornithorhynchus_anatinus	-----G	AGGTGAAGTC	GGTGAGATG	CACCACGAGG	CCCTAAGCGA	GGCCCTGCCA	GGAGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCCGTC-	[951]	
017_EF1a_Xenopus_laevis	-----G	AAGTAAAGTC	TGTGAAATG	CATCATGAGG	CCTTGCAAGA	AGCACTGCCT	GGAGACAATG	TGGGTTTCAA	TGTGAAGAAC	ATATCCGTC-	[951]	
018_EF1a_Saccharomyces_cerevisiae	-----G	AAGTCAAGTC	CGTTGAAATG	CATCAGGAAC	AATTGGAACA	AGGTGTTCCA	GGTGACAACG	TTGGTTTCAA	CGTCAAGAAC	GTTCCTGTC-	[945]	
019_EF1a_Neurospora_crassa	-----G	AAGTCAAGTC	CGTCGAGATG	CACCACGAGC	AGCTTGCTCTA	GGGTGTCCCT	GGTGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTTCCTGTC-	[948]	
020_EF1a_Cryptococcus_neoformans	-----G	AAGTCAAGTC	CGTTGAGATG	CACCACGAGC	AGATGCCCGA	GGGTCTTCCC	GGAGACAACG	TTGGTTTCAA	CGTCAAGAAC	GTTCCTGTC-	[945]	
021_EF1a_Schizosaccharomyces_mougeotii	-----G	AGGTCAAGTC	CGTCGAGATG	CACCACGAGC	AGCTCGCTGA	GGGCAAGCCG	GGTGACAACG	TCGGGTTCAA	CGTCAAGAAC	GTCTCCGTC-	[945]	
022_EF1a_Monoblepharis_insignis	-----G	AAGTTAAGTC	GGTTGAGATG	CACCACGAGT	CGCTCGCTGA	GGGTCTGCC	GGTGACAACG	TTGGTTTCAA	CGTGAAGAAC	GTGTCCGTC-	[831]	
023_EF1a_Rhizophyctis_harderi	-----G	AAGTCAAGTC	TGTGAAATG	CACCACGAAC	CCCTTGAAGC	CGGTCTCCCT	GGTGACAATG	TTGGATTCAA	CATCAAGAAC	GTTCCTGTC-	[831]	
024_EF1a_Harpochytrium_sp.	-----G	AAGTGAAGTC	CGTCGAGATG	CACCACGAGG	CCCTCCCCGA	GGCTCTCCCT	GGTGACAACG	TCGGGTTCAA	CGTCAAGAAC	GTCTCCGTC-	[783]	
025_EF1a_Absidia_glauca	-----G	AAGTCAAGTC	CGTCGAAATG	CATCAGGAAC	AATTGTTTGA	AGGTCTCCCT	GGTGACAACG	TCGGTTTCAA	CGTCAAGAAC	GTCTCCGTC-	[945]	
026_EF1a_Smittium_simulii	-----G	AAGTTAAATC	CGTAGAAATG	CATCAGGAAC	CACTCAGTGA	AGBTCTTCCA	GGAGACAACG	TTGGTTTCAA	CATCAAGAAC	GTTCCTGTC-	[918]	
027_EF1a_Chara_australis	-----G	AAGTCAAGTC	GGTGAGATG	CACCACGAGG	CGATGACCGA	GGCGTTCCG	GGCGACAACG	TGGGATTCAA	CGTGAAGAAC	GTAAAGCGTC-	[864]	
028_EF1a_Physcomitrella_patens	-----G	AGGTGAAGTC	GGTGAGATG	CATCAGAGT	CGATGCCTGA	GGCGACAACG	GGTGACAACG	TCGGGTTCAA	CGTGAAGAAC	GTAAAGCGTC-	[915]	
029_EF1a_Acetabularia_acetabulum	-----G	AGGTGAAGTC	GGTTGAGATG	CACCACGAAG	CTTTGCCCGA	AGCGGTGCCA	GGAGACAATG	TGGGTTTCAA	CGTGAAGAAC	GTGTCCGTC-	[864]	
030_EF1a_Chaetomorpha_coliformis	-----G	AGGTCAAAATC	GGTGAGATG	CACCACGAAG	CCCTTCCCGA	AGCTGTGCC	GGTGACAACG	TTGGGTTTAA	CGTGAAGAAC	GTAGTGTGTC-	[714]	
031_EF1a_Arabidopsis_thaliana	-----G	AGGTTAAGTC	TGTTGAGATG	CACCACGAGT	CTCTTCTTGA	GGCACTTCCC	GGTGACAATG	TTGGATTCAA	TGTCAAGAAC	GTTCCTGTC-	[915]	
032_EF1a_Oryza_sativa	-----G	AGGTCAAGTC	GGTTGAGATG	CACCACGAGG	CTCTGCCTGA	GGCTCTTCCC	GGTGACAACG	TTGGGTTTAA	CGTGAAGAAC	GTTCCTGTC-	[915]	
033_EF1a_Cyanidioschyzon_merolae	-----G	AAGTCAAGTC	AGTTGAAATG	CATCAGAGT	CACTCCCCGA	GGCTGGGCCA	GGTGACAACG	TTGGCTTCAA	CGTCAAGAAC	GTAGTGTCT-	[927]	
034_EF1a_Porphyrha_jezoensis	-----G	AGGTGAAGTC	GGTCGAGATG	CACCACGAGG	CGCTGCCCGA	GGCCGGCCCC	GGCGACAACG	TTGGCTTCAA	CGTCAAGAAC	GTGTCCGTC-	[927]	
035_EF1a_Plasmodium_falciparum	-----G	AATGTAATC	AGTTGAAATG	CACAAGGAAG	TTTGGGAAGA	AGCTAGACCA	GGAGACAACA	TTGGTTTTAA	CGTAAAGAAC	GTTCCTGTC-	[909]	
036_EF1a_Cryptosporidium_parvum	-----G	AAGTTAAGTC	AGTAGAAATG	CACCATGAGC	AGATGCCAGA	GGCCGTCCCA	GGTGACAACG	TTGGTTTCAA	TGTTAAGAAC	GTCTCCATC-	[903]	
037_EF1a_Tetrahymena_pyriformis	-----G	AAGTGAAGTC	CGTCGAAATG	CATCAGGAAC	AACCTCCCGA	AGCTGTCCCT	GGTGACAACG	TCGGTTTCAA	CGTAAAGGTC	GTCTCCGTC-	[918]	
038_EF1a_Stentor_coeruleus	-----G	AAGTTAAATC	AGTAGAAATG	CACCATGAAA	CCTTGAAGAT	TGCTAATCCA	GGGACAATG	TTGGATTCAA	CGTCAAGAAC	GTCTCAGTC-	[870]	
039_EF1a_Phytophthora_parasitica	-----G	AAGTCAAGTC	GGTGAGATG	CACCACGAGT	CTCTGCCTGA	GGCCGTCCCT	GGTGACAACG	TCGGCTTCAA	CGTCAAGAAC	GTTCCTGTC-	[915]	
040_EF1a_Mallomonas_rasilis	-----G	AAGTCAAGTC	TGTAGAGATG	CACCACGAGG	CCCTTGTCTGA	GGCTTCTCCA	GGTGATAACG	TTGGTTTCAA	CGTGAAGAAC	GTTCCTGTC-	[876]	
041_EF1a_Phaeodactylum_tricornutum	-----G	AAATTAAGTC	TGTCGAGATG	CACCACGAAT	CCCTTCCCGA	AGCGGTTTCT	GGTGACAATG	TTGGTTTCAA	CGTGAAGAAC	GTAGCCGTC-	[915]	
042_EF1a_Thalassiosira_pseudonana	-----G	ATGTGAAGAG	TATTGAGATG	CATCACAAC	AGATTCCAGA	GGCACTTCTC	GGTGACAATA	TCGGATTCAA	CATCAAGAAC	GTAGCAGTA-	[915]	
043_EF1a_Malawimonas_jakobiformis	-----G	AGGTCAAGTC	CGTTGAGATG	CATCAGGAAG	CGATGCCCGA	GGCCGTGCC	GGTGACAACG	TTGGCTTCAA	CGTGAAGAAC	GTTCGCCGTC-	[846]	
044_EF1a_Malawimonas_californiana	-----G	AGGTCAAGTC	CGTCGAGATG	CACCACGAGT	CGATGCCCGA	GGCCAGCC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTCTCCGTC-	[915]	
045_EF1a_Trimastix_pyriformis	-----G	AAGTCAAGTC	CGTCGAGATG	CACCACACCG	CTCTCGAGCA	GGCCACCCCT	GGTGACAACG	TCGGGTTCAA	CGTCAAGAAC	ATTGCCGTC-	[873]	
046_EF1a_Dinenympha_exilis	-----G	AGGTGAAGTC	AGTTGAGATG	CATCAGGAG	CTCTGACCCA	GGGACAATG	GGGACAATG	TGGGCTTCAA	CGTAAAGGTC	ATTGCTGTG-	[855]	
047_EF1a_Monocercomonoides_sp.	-----G	AAGTTAAATC	AGTTGAAATG	CATCATGAAG	CATTGGAACG	TGCTGTCCCA	GGTGACAACG	TTGGGTTTAA	TGTTAAGAAC	ATTTCAGTC-	[915]	
048_EF1a_Streblospio_tenuis	-----G	AGTGCAAAATC	AGTTGAGATG	CATCATGAAG	CTTTAACAAC	AGCCGTTCCTA	GGAGATAATG	TTGGATTTAA	CTTGAAGGTC	GTATCAGTC-	[864]	
049_EF1a_Histomonas_meleagridis	-----G	AAGTTAAGTC	CATCGAAATG	CACCACGAGC	AACCTCCAGA	GGCCGTCCCA	GGAGATAAAC	TCGGTTTCAA	CGTCAAGAAC	GTTCGCCGTC-	[915]	
050_EF1a_Trichomonas_vaginalis	-----G	AAGTTAAGTC	CATCGAAATG	CACCACGAAT	CCCTTCCAGA	GGCTTCTCCA	GGTGACAACA	TCGGCTTCAA	CGTCAAGAAC	GTTCGCCGTC-	[915]	
051_EF1a_Hexamita_inflata	-----G	AAGTTAAGTC	TGTCGAAATG	CACCACGAAC	AACCTCCAGA	GGCCGAGCCA	GGTGATAACG	TTGGTTTCAA	CATCAAGGTC	GTCTCCGTC-	[867]	
052_EF1a_Spiroplasma_vortens	-----G	AAGTCAAGTC	TGTCGAGATG	CACCACGAGG	AGCTCCCGCA	GGCCGAGCCA	GGCGACAACG	TCGGCTTCAA	CGTCCGTGTC	ATCGCTGCC-	[867]	
053_EF1a_Amalucyia_incarcerata	-----G	AAGTGAAGTC	CGTCGAAATG	CATCAGGAAG	CTCTTGAAGA	GGCTCTTCCA	GGTGACAACG	TCGGTTTCAA	CGTCAAGAAC	GTTCCTGTC-	[864]	
054_EF1a_Seculamonas_ecuadoriensis	-----G	AGGTCAAGTC	GATCGAGATG	CACCACACTC	AGCTTCCGGA	GGCCGTGCC	GGCGACAATG	TCGGTTTCAA	CATCAAGAAC	ATTGCCGTC-	[915]	
055_EF1a_Reclinomonas_americana	-----G	AGGTCAAGTC	CATTGAGATG	CACCACACCC	AGCTCCCGGA	GGCTTCCCT	GGAGACAACG	TCGGCTTCAA	CGTCAAGAAC	GTGTCCGTC-	[915]	
056_EF1a_Histiona_aroides	-----G	AAGTCAAGTC	GATCGAAATG	CACCACACCC	AGCTCCCGGA	AGCTGTGCCA	GGTGACAACG	TCGGCTTCAA	TGTCGAAGAAC	GTGTCCGTC-	[915]	
057_EF1a_Jakoba_libera	-----G	AGTGCAAAATC	CATCGAGATG	CACCACACTC	AGCTGCCCGA	GGCTCTGCC	GGTGACAACG	TTGGCTTCAA	CGTCAAGGTC	GTGTCCGTC-	[915]	
058_EF1a_Stachyamoeba_lipophora	-----G	ATGTCAAGTC	TGTTGAGATG	CATCAGAGC	AGCTCCCGGA	GGCTGTGCC	GGTGACAACG	TCGGCTTCAA	CGTCAAGAAC	ATCTCCGTC-	[912]	
059_EF1a_Naegleria_gruberi	-----G	AAGTCAAGTC	CGTCGAAATG	CATCAGACTT	CTGTTCCAGA	AGCTGTGCCA	GGTGACAATG	TTGGTTTCAA	CGTTAAGGTC	TTGTCCACC-	[867]	
060_EF1a_Naegleria_andersoni	-----G	AAGTCAAGTC	TGTCGAAATG	CATCAGACTT	CTGTCCAGGA	AGCTGTGCCA	GGTGACAACG	TTGGTTTCAA	CGTTAAGGTC	TTGTCCACC-	[885]	
061_EF1a_Acrasis_rosea	-----G	AAGTGAAGTC	TATCGAAATG	CATCAGGAAC	AACCTGTCTGA	GGCACTTCTC	GGAGACAACG	TTGGATTCAA	CATCAAGAAC	GTTCGCCGTC-	[864]	
062_EF1a_Sawyeria_marylandensis	-----G	AAGTTAAATC	TATTGAAATG	CATCATACTT	CAATTCTCTA	AGCAGAACA	GGAGATAACA	TTGGTTTTAA	CGTTAAAAAT	TTATCAGTC-	[915]	
063_EF1a_Bodo_saltans	-----G	AGGTGAAGTC	CGTCGAGATG	CACCACGAGT	CCCTCCCCGA	GGCCATCCCC	GGAGACAACG	TTGGCTTCAA	CGTGAAGAAC	TTGTCCATC-	[855]	
064_EF1a_Rhynchobodo_sp.	-----G	AAGTCAAGTC	CATCGAGATG	CACCACGAGT	CCCTCCCCGA	GGCCACCCCT	GGTGACAACG	TTGGTTTCAA	CGTCAAGAAC	ATCTCCGTC-	[849]	
065_EF1a_Rhynchomonas_nasuta	-----G	AGGTAAATC	AATCGAAATG	CACCACGAGA	GCATCCCCGA	GGCGATTCCC	GGAGACAACA	TCGGCTTCAA	CGTCAAGAAC	ATCTCCGTC-	[912]	
066_EF1a_Dimastigella_trypaniformis	-----G	AGGTCAAGTC	CATTGAGATG	CACCACGAGC	AGCTGGCTGA	GGCCGCCCGC	GGCGACAACG	TCGGCTTCAA	CGTCAAGAAC	ATCTCGATC-	[870]	

	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210	
067_EF1a_Perkinsiella_amoebae	-----	-G	AGGTCAAAGTC	CGTCGAGATG	CAYCACGAAG	CCCTCCCCGA	G6CCACGCCC	GGCATGAACA	TCGGCTTCAA	CGTCAAGAAC	GTCTCCGTC-	[840]
068_EF1a_Leishmania_major	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACGAGC	AGCTGGCGGA	G6CCGAAGCC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[915]
069_EF1a_Leishmania_infantum	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACGAGC	AGCTGGCGGA	G6CCGAAGCC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[915]
070_EF1a_Leishmania_braziliensis	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACGAGC	AGCTGGCGGA	G6CCGGTGCC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[915]
071_EF1a_Leishmania_tarentolae	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACGAGC	AGCTGGCGGA	G6CCGAAGCC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[915]
072_EF1a "L. tarentolae" FJ807257	-----	-G	AAGTGAAGTC	GATCGAAATG	CACCACGAGC	AGCTGGCGGA	G6CCGAAGCC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[840]
073_EF1a_Leptomonas_costaricensis	-----	-G	AGGTGAAGTC	GATCGAGATG	NATCGAGATG	-----	-----	-----	-----	-----	-----	[594]
074_EF1a_Leptomonas_bifurcata	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACGAGC	AGCTGGCTGA	G6CCGTG6CC	GGCGACAACG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[675]
075_EF1a_Leptomonas_podlipaevi	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACGAGC	AGCTGGCCGA	G6CCGTG6CC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTCTCCGTG-	[663]
076_EF1a_Herpetomonas_muscarum	-----	-G	AGGTCAAAGTC	CATCGAGATG	CACCACGAGC	AGCTCGCTGA	G6CCGTCCCC	GGCGACAACG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[828]
077_EF1a_Herpetomonas_pestoi	-----	-G	AGGTCAAAGTC	CATCGAGATG	CACCACGAGC	AGCTGGCTGA	G6CCGTG6CC	GGTGAACAAC	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[696]
078_EF1a_Blastocrithidia_culicis	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACG	-----	-----	-----	-----	-----	-----	[766]
079_EF1a_Trypanosoma_brucei	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACGAGC	AGCTCGCTGA	G6CCACCCCC	GGTGAACAAC	TCGGCTTCAA	CGTGAAGAAC	GTCTCTGTA-	[915]
080_EF1a_T. brucei_equiperdum	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACG	-----	-----	-----	-----	-----	-----	[774]
081_EF1a_Trypanosoma_avium	-----	-G	AGGTGAAGTC	AATTGAGATG	CACCACGAGC	AGCTGGCGGA	G6CCACCCCT	GGTGAACAAC	TGGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[861]
082_EF1a_Trypanosoma_cruzi	-----	-G	AGGTGAAGTC	GATTGAGATG	CACCACGAGC	AGCTGGCCGA	G6CCGACGCC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[915]
083_EF1a_Trypanosoma_rangeli	-----	-G	AGGTGAAGTC	GATCGAGATG	CACCACGAAC	AGCTGGCCGA	G6CCACGCCC	GGTGAACAAC	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[915]
084_EF1a_Ploetia_costatum	-----	-G	AAGTGAAGTC	TGTTGAGATG	CACCACGAGG	C6CTGGCTGA	G6CCATCCCC	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
085_EF1a_Ploetia_oblonga_cDNA	-----	-G	AAGTGAAGTC	TGTTGAGATG	CACCACGAGG	C6CTGGCTGA	G6CCATCCCC	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
086_EF1a_Ploetia_oblonga_gDNA	-----	-G	AGGTCAAAGTC	CATTGAGATG	CATCACGAGT	CCCTCCCCGA	G6CCGTG6CC	GGCGACAACG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
087_EF1a_Entosiphon_sp.	-----	-G	AGGTGAAGTC	CATTGAGATG	CACCACGAGG	C6CTGGCTGA	G6CCGTG6CC	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
088_EF1a_Entosiphon_sulcatum	-----	-G	AGGTGAAGTC	CATTGAGATG	CACCACGAGG	CCCTGCCAGA	G6CCGTG6CC	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[840]
089_EF1a_Peranema_tricophorum	-----	-G	AAGTTAAAGTC	CATCGAAATG	CACCACGAAG	CTCTTGCCGA	AGCCATCCCT	GGTGAACAAC	TTGGATTCAA	CGTGAAGAAC	GTCTCCGTG-	[912]
090_EF1a_P. tricophorum_FJ807243	-----	-G	AAGTTAAAGTC	CATCGAAATG	CACCACGAAG	CTCTTGCCGA	AGCCATCCCT	GGTGAACAAC	TTGGATTCAA	CGTGAAGAAC	GTCTCCGTG-	[906]
091_EF1a_Colacium_elongatum	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGT	CTCTTGCCGA	G6CTGTCCCT	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
092_EF1a_Colacium_mucronatum	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGT	CCCTGGCTGA	G6CCATCCCC	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[858]
093_EF1a_Colacium_sp.	-----	-G	AGGTGAAGTC	C6TTGAGATG	CACCACGAGT	CTCTTGCCGA	G6CTATCCCT	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
094_EF1a_Colacium_vesiculosum	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGT	CCCTGGCTGA	G6CCATCCCC	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[870]
095_EF1a_Cryptoglena_pigra	-----	-G	AGGTGAAGTC	AGTTGAGATG	CACCACGAGG	CTCAGGTTGA	G6CTGTCCCT	GGCGACAACG	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
096_EF1a_Cryptoglena_skujai	-----	-G	AGGTGAAGTC	CGTTGAGATG	CACCACGAGG	CTTTGGCAGA	G6CCGTG6CC	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
097_EF1a_Discoplastis_adunca	-----	-G	AGGTGAAGTC	TGTCGAAATG	CATCACGAGT	CTCTCACTGA	G6CCATCCCT	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
098_EF1a_Euglena_clara	-----	-G	AGGTGAAGTC	GGTCGAGATG	CACCACGAGT	G6CTGTG6GA	G6CCATCCCT	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[792]
099_EF1a_Euglena_deses	-----	-G	AGGTGAAGTC	AGTCGAGATG	CACCACGAGT	CTCTTCTTGA	G6CCATCCCT	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
100_EF1a_Euglena_exilis	-----	-G	AAGTCAAAGTC	AGTGGAAATG	CATCACGAAG	CTCTGGCTGA	AGCTATTCCA	GGAGACAATG	TTGGATTCAA	CGTGAAGAAC	GTATCTGTC-	[792]
101_EF1a_Euglena_geniculata	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAGG	C6CTGGCTGA	G6CCATCCCC	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
102_EF1a_Euglena_gracilis	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAGG	CTCTGACCGA	G6CCATCCCT	GGCGACAACG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
103_EF1a_E. gracilis_X16890	-----	-G	AAGTCAAAGTC	AGTTGAGATG	CACCATGAGG	CTCTGACCGA	AGCTGTG6CT	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[915]
104_EF1a_E. gracilis_FJ807238	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAGG	CTCTGACCGA	G6CCATCCCT	GGCGACAACG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[915]
105_EF1a_Euglena_granulata	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAGG	C6CTGCTTGA	G6CCATCCCT	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
106_EF1a_Euglena_hiemalis	-----	-G	AAGTCAAAGTC	CGTGGAGATG	CACCACGAGG	CTCTCACTGA	G6CTATTCCT	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
107_EF1a_Euglena_longa	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAGG	CTCTGACCGA	G6CCATCCCT	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[429]
108_EF1a_Euglena_stellata	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAGT	CTCTTCTGGA	AGCCATTCCA	GGAGACAATG	TGGGCTTCAA	CGTGAAGAAC	GTGTCAGTA-	[792]
109_EF1a_Euglena_viridis	-----	-G	AAGTCAAAGTC	C6TTGAGATG	CATCACGAGT	CTCTGGCTGA	R6CTATCCCT	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[870]
110_EF1a_Eutreptia_pertyi	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAGG	C6CTGGCTGA	G6CCATCCCT	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
111_EF1a_Eutreptiella_pomquetensis	-----	-G	AAGTGAAGTC	TGTTGAGATG	CACCACGAGG	C6CTGGCTGA	G6CCATCCCC	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[522]
112_EF1a_Lepocinlis_acus_var._major	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGT	CTCTTGACGA	G6CGTCCCT	GGCGACAACG	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
113_EF1a_Lepocinlis_ovum	-----	-G	AGGTGAAGTC	AGTTGAGATG	CACCACGAGT	C6CTG6CAGA	AGCTGTCCCT	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[792]
114_EF1a_Lepocinlis_spyrogyroides_clone_38	-----	-G	AAGTGAAGTC	TGTTGAGATG	CACCACGAGT	CTCTGGCTGA	AGCTATTCCC	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[225]
115_EF1a_L. spyrogyroides_clones_1_&_4	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGG	CTCTCACTGA	G6CCGTCCCT	GGGTTCAATG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
116_EF1a_Lepocinlis_steinii	-----	-G	AAGTGAAGTC	C6TTGAGATG	CACCACGAGT	C6TTGGCAGA	G6CCATCCCC	GGCGACAACG	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[792]
117_EF1a_Lepocinlis_triapteris	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGG	TGTTGGAGATG	G6CTATCCCA	GGTGAACAAC	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
118_EF1a_Monomorphina_aenigmatica	-----	-G	AGGTGAAGTC	AGTTGAAATG	CACCACGAGT	C6TTGGCTGA	G6CCATCCCT	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
119_EF1a_Monomorphina_pyrum	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGG	CTCTCACTGA	G6CCGTCCCT	GGGACAATG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
120_EF1a_Monomorphina_sp.	-----	-G	AAGTGAAGTC	CGTCGAGATG	CACCACGAGG	CTCTGACTGA	G6CTGTG6CT	GGTGATAATG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
121_EF1a_Phacus_orbicularis	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAAG	CCTCACTGA	G6CCATTCCT	GGGACAACG	TTGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[870]
122_EF1a_Phacus_similis	-----	-G	AAGTCAAAGTC	CGTCGAGATG	CAYCACGAAG	CCCTCCCGGA	G6CCATCCCT	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
123_EF1a_Strombomonas_acuminata	-----	-G	AAGTGAAGTC	TGTTGAAATG	CATCACGAGT	C6CTGGCAGA	G6CTGTG6CT	GGTGATAACG	TTGGTTTCAA	TGTTGAAGAAC	GTATCCGTG-	[915]
124_EF1a_Strombomonas_borsteniensis	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGT	CCCTGGCTGA	G6CTGTG6CT	GGTGAACAAC	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[792]
125_EF1a_Strombomonas_napiformis	-----	-G	AGGTGAAGTC	TGTTGAAATG	CACCACGAGT	C6TTGGCTGA	G6CTGTG6CT	GGTGAACAAC	TATGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
126_EF1a_Strombomonas_ovalis	-----	-G	AGGTGAAGTC	CGTTGAGATG	CACCACGAGT	CCCTGGCTGA	G6CTGTG6CT	GGTGAACAAC	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
127_EF1a_Strombomonas_triquetra	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCACGAGG	CTCTGGCTGA	G6CTGTG6CT	GGGACAATG	TCGGTTTCAA	-----	-----	[891]
128_EF1a_Trachelomonas_conradi	-----	-G	AAGTCAAAGTC	CGTCGAGATG	CACCACGAGT	CTCTGGCTGA	G6CCATCCCC	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
129_EF1a_Trachelomonas_echinata	-----	-G	AGGTGAAGTC	TGTTGAGATG	CACCATGAGT	CTTTGGCCGA	G6CCATCCCA	GGTGAACAAC	TTGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
130_EF1a_Trachelomonas_lefevrei	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAAG	CCTTGACTGA	G6CTATTCT	GGTGAACAAC	TGGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[912]
131_EF1a_Trachelomonas_plantonica	-----	-G	AAGTCAAAGTC	TGTTGAGATG	CACCACGAAG	CCTTGACCGA	G6CTATTCT	GGAGACAACG	TGGGTTTCAA	CGTGAAGAAC	GTGTCGGTG-	[792]
132_EF1a_Trachelomonas_similis	-----	-G	AGGTCAAAGTC	TGTTGAAATG	CACCACGAAT	CTTTGCTTGA	G6CTGTG6CT	GGAGACAACG	TTGGTTTCAA	CGTGAAGAAC	GTGTCAGTG-	[792]

		1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210		
133	EFL1a	Trachelomonas_volvocina	-----G	AAGTCAAGTC	CGTCGAGATG	CACCACGAGG	CATTGTGCA	GGCTGTCCCT	GGTGACAACG	TTGGTTTCAA	CGTGAAGAAC	GTGTCCGCTC	[912]	
134	EFL	environmental_sample_EU810232	CGCC-----	GGAA AGGTGTTCTC	CATCGAGATG	CACCACAAGA	CCCTCGACGC	CGCCCTCCCG	GGACACAACG	TCGGCATGAA	CATCAAGGGA	CTGCCCAAGG	[1003]	
135	EFL	Creolimax_fragrantissima	CTCT-----	GGAA AGGTTTTTAC	CGTCGAGATG	CACCACAAGC	AGGTGCGAGC	CGCCTACGCT	GGAGACAACG	TTGGTCTTAA	CATCAAGGGT	CTTAACGCCA	[931]	
136	EFL	Sphaeroforma_arctica	CACC-----	GGAA AGGTGTTTAC	CGTCGAGATG	CACCACAAGC	AGGTTGAGGC	CGCTTACGCT	GGTGACAATG	TTGGTCTTAG	CATCAAGGGT	CTTAACGCCA	[1003]	
137	EFL	Monosiga_brevicollis	CACC-----	GGCA AGGTCTTACG	CGTCGAGATG	CACCACAAGT	CGCTGAGACT	TGCCATGACT	GGTGACAACG	TTGGCCTCAA	CATCAAGGGT	CTCAACAAGG	[1003]	
138	EFL	Xenopus_laevis	CACC-----	GGCA AGGTGTTTCT	TGTTGAGATG	CACCACAAGC	AGGTGCGAGT	GGCCGGCCCT	GGCGACAACG	TCGGCATGAA	CGTCAAGAAC	TTGAACGCCA	[1003]	
139	EFL	Spizellomyces_punctatus	TGGT-----	GGAA AGGTCTTAC	CGTCGAAATG	CACCACAAGA	AGGTGGAGAA	GGGTCTTGCC	GGTGACAACA	TCGGTATGAA	CATCAAGGGT	CTTGACAAGA	[1009]	
140	EFL	Gaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]	
141	EFL	Allomyces_macrognus	CGCC-----	GGCA AGGTCTTCTC	GGTCGAGATG	CACCACAAGA	AGGTCGAGAA	GGCGGAGACG	GGCGACAACG	TCGGCATGAA	CATCAAGGGT	CTCGACAAGA	[1009]	
142	EFL	Blastocladiella_emersonii	CGCT-----	GGTA AGGTGTTTAC	GGTCGAGATG	CACCACAAGA	AGGTGCGAGAA	GGCGGAGACT	GGCGACAACG	TCGGCATGAA	CATCAAGGGT	CTCGACAAGA	[1009]	
143	EFL	Catenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]	
144	EFL	Coelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]	
145	EFL	Entomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]	
146	EFL	Basidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]	
147	EFL	Conidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]	
148	EFL	Tetraselmis_tetrathele	CACC-----	GGCA AAGTGTTCAC	GGTGGAGATG	CACCACAAGC	GGCTGGAGGC	CGCTGGCCCC	GGCGACAACG	TCGGCATGAA	CATCAAGGGT	CTTGAGAAGA	[952]	
149	EFL	Mesostigma_viride	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[879]	
150	EFL	Ostreococcus_tauri	CGGT-----	GGTA AGGTGTTTAC	CGTCGAAATG	CACCACAAGC	GTGTCGAAGC	GGCCGCGCCG	GGTGACAACG	TCGGTATGAA	CATCAAGGGT	CTCGACAAGC	[1012]	
151	EFL	Nephroselmis_olivacea	CGTC-----	GGCA AGGTGTTTCT	CGTGGAGATG	CACCACAAGT	CGGTGGCCGA	GGCCAAGGCC	GGTGACAACG	TCGGCATG--	-----	-----	[957]	
152	EFL	Micromonas_pusilla	CGTC-----	GGCA AGGTGTTTCT	CGTCGAGATG	CACCACAAGC	CGCTCGAGGA	GGCTCACCCC	GGTGACAACG	TCGGCATGAA	CATCAAGGGT	CTCGACAAGA	[961]	
153	EFL	Parachlorella_kessleri	CATC-----	GGCA AGGTCTTAC	CGTGGAGATG	CACCACAAGC	GGCTGGACAA	GGCCGGACCC	GGTGACAACG	TCGGCATG--	-----	-----	[942]	
154	EFL	Helicosporidium_sp.	CACC-----	GGCA AGGTGTTTAC	CATCGAGATG	CACCACAAGC	GGCAGGAGCA	GGCCGGCCCC	GGCGACAACG	TTGGCATGAA	CGTCAAGGGT	CTCGACAAGA	[1024]	
155	EFL	Gonium_pectorale	CTCC-----	GGCA AGGTGTTTCT	CGTGGAGATG	CACCACAAGC	GGTGGGACAA	GGCCGGCCCC	GGCGACAACG	TCGGCATGAA	CATCAAGGGT	CTGGACAAGG	[1006]	
156	EFL	Chlamydomonas_reinhardtii	CACC-----	GGCA AGGTGTTTAC	CGTGGAGATG	CACCACAAGC	GTGTTGGACAA	GGCCGGCCCC	GGCGACAACG	TCGGCATGAA	CATCAAGGGT	CTGGACAAGG	[1006]	
157	EFL	Scenedesmus_obliquus	CGTG-----	GGCA AGGTGTTTCT	CGTGGAGATG	CACCACAAGC	GGTGGGACAA	GGCCGGCCCC	GGTGACAACG	TCGGCATGAA	CATCAAGGGC	CTGGACAAGG	[1006]	
158	EFL	"Oryza_sativa"	CGTG-----	GGCA AGGTGTTTCT	CGTGGAGATG	CACCACAAGC	GTGTTGGACAA	GGCCGGCCCC	GGTGACAACG	TCGGCATGAA	CATCAAGGGC	CTGGACAAGG	[1006]	
159	EFL	Chlorococcum_sp.	CGAG-----	GGCA AGGTGTTTCT	TGTGGAGATG	CACCATAAGC	GTGTCGACCA	GGCTGGCCCC	GGTGACAACG	TTGGCATGAA	CATCAAGGGT	CTTGACAAGG	[955]	
160	EFL	Ochlochaete_hystrix	CACT-----	GGCA AGGTGTTTCT	CGTGGAGATG	CACCACAAGC	GGTGGGACAA	GGCCGGCCCC	GGTGACAACG	TGGGCATGAA	CATCAAGGGT	CTGAACAAGG	[964]	
161	EFL	Ulva_intestinalis	TGAG-----	GGCA AGGTGTTTCT	TGTGGAGATG	CACCACAAGC	GTGTTGACTC	GGCCGGCCCC	GGTGACAATG	TCGGCATGAA	CATCAAGGGC	TTGGACAAGG	[955]	
162	EFL	Urospora_sp.	CGAG-----	GGCA AGGTGTTTCT	CGTGGAGATG	CACCACAAGC	GGCTCGAGCA	GGCCGGCCCC	GGCGACAACG	TCGGCATGAA	CATCAAGGGC	CTGCACAAGG	[808]	
163	EFL	Gymnochlorella_stellata_ppc	TAAC-----	GGAA AAGTCTTCTC	TATTGAGATG	CATCATAAGC	GAATGATGGA	AGCTCTGCTC	GGTGACAACG	TCGGAATGAA	CGTTAAGGGC	TTCCCAAGG	[1042]	
164	EFL	Lotharella_amoebiformis_ppc	CGCC-----	GGAA AGGTGTTTCT	CATCGAGATG	CATCATAAGC	GATTTCAGAA	GGCCTATGCG	GGTGATAACG	TCGGTATGAA	CGTGAAGGGG	TTCCCAAGG	[787]	
165	EFL	Chlorarachnion_reptans_ppc	CTCC-----	GGAA AAGTGTTCCT	TATCGAAATG	CATCACAAGC	GTATTGAAC	TGCCAAGCCA	GGAGACAATG	TAGGAATGAA	CGTGGCTGGT	TTCCCTAAGG	[814]	
166	EFL	Chondrus_crispus	TTCT-----	GGAA AGGTGTTTCT	CGTCGAAATG	CATCACAAGT	CAGTGCCGGA	GGCTGTGCCC	GGAGACAATG	TCGGCATGAA	TGTCAAGGGT	GTGCGAAGG	[1003]	
167	EFL	Gracilaria_changii	TTCT-----	GGTA AGGTGTTTCT	AGTCGAAATG	CATCATAAGT	GTGTTGAAGA	GGCTGGGACA	GGCGACAACG	TCGGTCTCAA	TATCAAGGGT	CTCGAGAAGT	[1003]	
168	EFL	Cryptomonas_ovata	CGCG-----	GGCA AGATCTTAC	CGTGGAGATG	CACCACAAGC	GTATCGATGC	GGCCAAGCCT	GGTGACAACG	TGGGCATGAA	CATCAAGGGT	CTGGACAAGA	[961]	
169	EFL	Guillardia_theta	CGAA-----	GGCA AGATCTTAC	TGTTGAGATG	CACCACAAGC	GTGTTGACGC	GGCCAAGCCT	GGTGACAACG	TCGGCATGAA	CATCAAGGGT	CTTGACAAGA	[1006]	
170	EFL	Rhodomonas_salina	CGCC-----	GGCA AGATCTTAC	CGTGGAGATG	CACCACAAGC	GTATCGACGC	GGCGAAGCCC	GGCGACAACG	TTGGCATGAA	CATCAAGGGT	CTCGACAAGA	[1003]	
171	EFL	Pavlova_lutheri	CGTG-----	GGCA AGGTGTTTCT	CGTCGAGATG	CACCACAGCG	GCATCGACCA	GGCCGGCCCC	GGTGACAACG	TCGGCATGAA	CATTAAGGGC	CTGCTCAAGG	[1003]	
172	EFL	Chrysochromulina_sp.	CGGC-----	GGCA AGGTCTTAC	CATTGAGATG	CACCACACCC	GCCACGAGGT	GGCCAAGCCC	GGCGACAACG	TCGGCTCAA	CATCAAGGGT	CTCGACAAGA	[823]	
173	EFL	Isochrysis_galbana	CGCG-----	GGCA AGGTATTAC	CATCGAGATG	CACCACAGCG	GGCAGGAGT	TGCCAAGCCC	GGCGACAACG	TCGGCTCAA	CATCAAGGGC	CTCGACAAGA	[1006]	
174	EFL	Emiliana_huxleyi	CGCG-----	GGCA AGGTCTTAC	CATCGAGATG	CACCACAAGA	GGCACCCGCA	AGCCAACCCC	GGCGACAACG	TCGGCTCAA	CATCAAGGGC	CTCGAGAAGC	[1012]	
175	EFL	Achnanthes_kuwaitensis	-----	AAGA AGATCTTCTC	CATTGAGCAG	CACAAGAAGA	CTCTCGATTG	TGCCGAGCCC	GGAAACTCTG	TCGGCTTTC	CATTAAGGGC	ATTGCCAAGG	[913]	
176	EFL	Asterionella_glacialis	-----	AAGA AGATGTTTCT	CATCGAACAA	CACAAGAAGG	TCCTCGAATC	GGCCGGTCCC	GGTAACCTCT	TTGGTATGTC	CATCAAGGGT	ATCACCAAGG	[913]	
177	EFL	Detonula_confervacea	-----	AAGA AGATCTTCTC	CATCGAACAG	CACAAGAAGC	AGCTCGAGAG	TGCCACCCCT	GGAGACTCCG	TTGGTCTTCT	CATCAAGGGT	ATTGCCAAGG	[913]	
178	EFL	Thalassiosira_pseudonana	-----	AAGA AGGTCTTCTC	CATCGAACAA	CACAAGAAGC	AACCTGAGAA	GGCTTGCCCC	GGAGACTCTG	TCGGTCTTTC	CATCAAGGGT	ATTGCCAAGG	[976]	
179	EFL	Perkinsus_marinus	TACT-----	GGTA AGGTCTTTAC	TGTTGAGATG	CATCACACCC	GTACAGGAGCA	GGCTCTTCCG	GGTGACAACG	TTGGTCTGAA	CGTTAAGGGT	TTGCCGAAGG	[1021]	
180	EFL	Oxyrrhis_marina	CTCT-----	GGGA AGGTGTTTAC	TGTGGAGATG	CACCACACTC	GTACGAGAGCA	GGCCAAGCCC	GGTGACAACG	TTGGGTGAA	CATCAAGGGT	CTGGAGAAGG	[1015]	
181	EFL	Karlodinium_micrum	CACA-----	GGCA AGGTGTTTAC	TGTGGAGATG	CATCACTCCC	GAGTGGATGC	TGCTAACCTT	GGAGACAACG	TTGGGCTAAA	CATCAAGGGC	TTGGATAAGA	[1009]	
182	EFL	Amphidinium_carterae	CACA-----	GGCA AGGTGTTTAC	AGTGGAGATG	CACCACACTC	GCATGGACCA	AGCCAACCCA	GGTGACAACG	TTGGCTCAA	CATCAAGGGT	CTTGACAAGA	[1009]	
183	EFL	Heterocapsa_triquetra	CACG-----	GGCA AGGTCTTAC	CGTGGAGATG	CACCACCTGC	GGCTGGACTT	GGCGAAGCCC	GGTGACAACG	TTGGCTGAA	CATCAAGGGT	CTTGACAAGA	[1009]	
184	EFL	Planoglabratella_operculis	-----	GGAA AAGCGTTTTT	CATTGAGATG	CATCATAAAA	ATGTCGATAA	AGCGGAATGT	GGTGACAATG	TTGGTGTCAA	TGTTAAGGGA	TTAAACAAGG	[829]	
185	EFL	Reticulomyxa_filosa	-----	GGCA AAGCATTTTT	TATTGAAATG	CACCACAAGA	AGTGGACAA	CAGTGAAGCT	GGTGACAACG	TCGGTGTAAA	CGTGAAGAAC	TTGAAAAAGG	[904]	
186	EFL	Lotharella_amoebiformis	CCCTTCGCGG	CTTGAGGCTA	AGGTCTTCTC	CATCGAGGCT	CACCACGTC	AGCTTCCCAA	GGCTGAGGCT	GGTGATAACA	TTGGAATTTG	TATTAAGGGT	CTCCCAAGG	[1084]
187	EFL	Gymnochlorella_stellata	CCCATCTGGA	CTTAGCTGCA	AGGTCTTCTC	TATCGAAGCT	CACCACAGAC	AGCAAGCTAG	GGCTTCCGCT	GGGGAACAAC	TCGGTATCTG	CATTAAGAAC	ATGAAGAAGG	[1042]
188	EFL	Chlorarachnion_reptans	CCCATCCGGT	CTCAGATGCA	AGGTCTTCTC	CATCGAAGCT	CACCACAGAC	AGCAGCGTAG	GGCTGTGCTG	GGGTATAACG	TCGGTATTTG	TATCAAGGGG	CTTCTCAAGA	[973]
189	EFL	Diplonema_ambulator	CACC-----	GGGA AGGTGTTTAC	CTCGAGATG	CACCACAAGA	GGTGGCAGCA	CGCGAAGGCC	GGGGAACAAC	TCGGGCTGAA	CATCAAGGGT	CTGAACAAGG	[607]	
190	EFL	Diplonema_papillatum	CACT-----	GGTA AGGTGTTTAC	GTGCGAGATG	CACCACAAGA	GGTGTGACGA	CGCGAAGGCC	GGAGACAACG	TTGGTTTGA	CGTGAAGGCC	CTGAACAAGG	[979]	
191	EFL	Rhynchopus_euleeides	CACC-----	GGCA AGGTGTTTCT	GTGCGAGATG	CACCACAAGC	GTGCGACGA	CGCCAAGGCC	GGTGACAACG	TCGGCTGAA	CGTGAAGGCC	CTGAACAAGG	[622]	
192	EFL	Petalomonas_cantuscygni	CTCC-----	GGCA AGGTGTTTCT	CATCGAGATG	CATCACAGC	GATTGAGAA	GGCCATGCCG	GGCGACAACG	TCGGCTGAA	CATCAAGGGT	CTCGACAAGC	[850]	
193	EFL	Trypanoplasma_borrel	CCAG-----	GGTA AGGTGTTTCT	CATTGAAATG	CACCACAAGG	AGCATGCCAC	CGCGGAGGAC	GGCGACAACG	TCGGCTCAA	CATCAAG--	-----	[1017]	
194	EFL	Neobodo_salien	CGTC-----	GGTA AGGTGTTTCT	GATCGAGATG	CACCACCGCG	AGCACGACAA	GGCCGGCCCC	GGCGACAACG	TCGGCTGTC	GATGAAGGGT	CTGCCAAGG	[976]	

	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	
001_aEF1a_Desulfurococcus_mobilis	--AAGGATAT	TAAGCGTGGG	GACGTAGCAG	GCTCACTGGA	T-----	GTC-----	C CA--CCAAC	CGTGGCA---	-----	G ACGAGTTAC	GGCTAGAATA	[996]
002_aEF1a_Sulfolobus_solfataricus	--AAGATAT	AAAAAGAGGA	GATGTTGTAG	GTCACTCCGAA	T-----	AAT-----	C CA--CCTAC	TGTTGCT---	-----	G ATGAATTAC	TGCAAGAATA	[990]
003_aEF1a_Thermoplasma_acidophilum	--aAcGaTaT	CaaGaGgGgA	gAtGtctGcG	gTcAcctCgA	t-----	aCa-----	C cC--CcGaC	cGtTgTg---	-----	A aGgCgTtCaC	cGtCagAtA	[972]
004_aEF1a_Halobacterium_sp.	--GAGGACAT	TCGTCGCGGC	GACGCTCTGT	GTCCGGCCGA	C-----	GAC-----	C Cg--CCGTC	GGTCGCC---	-----	G ACACCTTCA	AGCGCAGGT	[963]
005_aEF1a_Methanocaldococcus_jannaschii	--AAGATAT	TAAGAGAGGA	GACGTTTGTG	GGACCCGAGA	C-----	AAC-----	C CA--CCAAC	AGTTGCA---	-----	G AAGAATTAC	AGCTCAAATC	[978]
006_EF1a_Acanthamoeba_culbertsoni	--AAGGACAT	CGCGCTGTGT	AACGTGTGCG	GTGACTCCAA	G-----	AAG--GACC	Cg--CCCCA	GGAGACT---	-----	G AGGACTTAC	CGCCACAGGT	[963]
007_EF1a_Neoparamoeba_branchipila	--AAGGACAT	CCGTCGTGGT	AACGTTGCTG	GTGACTCCAA	G-----	AAC--GACC	CC--CCCGC	TGGTGCC---	-----	G AGGACTTAC	TGCCCAAGTT	[948]
008_EF1a_Planoprotostelium_aurantium	--AAGGACAT	CGCGAGAGGT	TACGTCGCTG	GTGACGCCAA	G-----	GCC--GACC	CC--CTCTC	TGGAAGT---	-----	G AGTTCTTCG	CGCTCAGGTT	[945]
009_EF1a_Physarum_polycephalum	--AAGGATAT	CGCGCTGGT	ATGCTGCGTG	GTGACTCCAA	G-----	AAC--AGCC	CA--CCCGC	TAGAGCT---	-----	G AGTCTTCAA	CGCCCAAGTT	[945]
010_EF1a_Amoebidium_parasiticum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[867]
011_EF1a_Capsaspora_owczaraki	--AAGGACAT	CGCTCGTGGC	AACGTTGCGG	GGGACTCGAA	G-----	AAC--GACC	CC--CCGAA	GGAGACC---	-----	A AGACCTTAC	CGCCACAGGT	[1026]
012_EF1a_Monosiga_ovata	--AAGGACAT	TCGCCGTGGC	AACGTTTGGC	GTGACTCCAA	G-----	AAC--GACC	CC--CCGAA	GGCTTCT---	-----	G AGTCTTCTC	GGCTCAGGT	[1005]
013_EF1a_Salpingoeca_amphoridium	--AAGGATAT	CGCGCTGGC	AACGTCGCGC	GTGACTCGAA	G-----	AAC--GACC	CT--CCCAA	GGCCACT---	-----	G AGAAGTTCA	CGCCACAGGT	[1008]
014_EF1a_Ceanorhabditis_elegans	--AAGGATAT	TCGTCGTGGA	TCCGTCGCTC	CGGACTCCAA	G-----	CAA--GACC	CA--GCCAA	GGAGGCT---	-----	C GCACCTTCA	CGCCACAGGT	[1032]
015_EF1a_Drosophila_melanogaster	--AAGGAGCT	CGCTCGTGGC	TACGTTGCGC	GTGACTCCAA	G-----	GCT--AACC	CC--CCCAA	GGGAGCC---	-----	G CGGACTTAC	CGCCACAGGT	[1032]
016_EF1a_Ornithorhynchus_anatinus	--AAGGACAT	CGCCCGGGGC	AACGTCGCGC	GAGACAGCAA	G-----	TCG--GACC	CA--CCCAA	GGAGACCA---	-----	G CTCAGTTAC	TTCCCAAGGT	[1032]
017_EF1a_Xenopus_laevis	--AAGGACAT	TAGGAGAGGA	AATGTAGCTG	GGGATAGCAA	A-----	AAT--GACC	CA--CCTAT	GCAAGCT---	-----	G GCAGCTTAC	AGCACAGGT	[1032]
018_EF1a_Saccharomyces_cerevisiae	--AAGGAAT	CAGAAGAGGT	AACGTCGTGT	GTGACGCTAA	G-----	AAC--GATC	CA--CCAAA	GGGTTGC---	-----	G CTCTTTTCA	CGCTACCGTC	[1026]
019_EF1a_Neurospora_crassa	--AAGGATAT	CGCCGCTGGT	AACGTTGCGC	GTGACTCCAA	G-----	AAC--GACC	CC--CTGTC	TGGCGCC---	-----	G CTTCTTTCAC	CGCCACAGGT	[1029]
020_EF1a_Cryptococcus_neoformans	--AAGGACAT	CGCAGGAGGT	AACGTCGTGT	GTGACTCCAA	G-----	AAC--GACC	CC--CCTAT	GGAGGCT---	-----	G CTTCTTTCAC	CGCCACAGGT	[1026]
021_EF1a_Schizophyllum_commune	--AAGGACAT	TCGCCGTGGT	AACGTCGCTC	CGGACTCCAA	G-----	AAC--GACC	CC--CGCAA	GGAGGCG---	-----	G CGTCTTTCAC	CGCCACAGGT	[1026]
022_EF1a_Monoblepharis_insignis	--AAGGACAT	TCGCCGTGGT	AACGTTGCGT	CGGACTCGAA	G-----	AAC--GACC	CC--GCTCA	GGAGTCC---	-----	G CCTCGTTCA	CGCTCAGGT	[912]
023_EF1a_Rhizophyctis_harderi	--AAGGAAT	CGCTCGTGGT	ATGTTTGTGT	CTGACACCAA	G-----	AAC--GACC	CT--GCCAA	GGAGGCT---	-----	G CCTCATTCAC	TGCTCAAATT	[912]
024_EF1a_Harpochytrium_sp.	--AAGGATAT	CGCTCGTGGG	AACGTCGCGC	GAGACTCCAA	G-----	AAC--GACC	CC--CCCAA	GGAAACC---	-----	G CCTCGTTCA	CGCCACAGGT	[864]
025_EF1a_Absidia_glauca	--AAGGATAT	CGCTCGTGGT	AACGTCGTGT	CTGACTCCAA	G-----	AAC--GACC	CC--GCCAA	GGAAAGCT---	-----	G GCTCCTTAC	TGCCCAAGTC	[1026]
026_EF1a_Smittium_simulii	--AAGATAT	TAGACTGTGT	AACGTTTGGC	GTGACTCCAA	G-----	AAT--GATC	CA--CCAAA	GGAAACT---	-----	G GAACCTTAC	TGCTCAAGTT	[999]
027_EF1a_Chara_australis	--AAGGAGTT	GAGGAGAGGT	TTGTTGGCTC	CGGACTCTAA	G-----	AAC--GACC	CT--GCCAA	GGAGGCG---	-----	G CGAGCTTAC	GTCTCAGGT	[945]
028_EF1a_Physcomitrella_patens	--AAGGACCT	GAGGAGAGGT	TACGTTGGCT	CAGACTCGAA	G-----	AAC--GACC	CT--GCCAA	GGAGGCA---	-----	G CCAACTTAC	CGCCACAGGT	[996]
029_EF1a_Acetabularia_acetabulum	--AAGGAGAT	CAAGCGGGGG	TATGTGGCAT	CAGACTCCAA	A-----	AGT--GATC	CC--GCTAA	AGGCTGC---	-----	G AGAAGTTAC	GGCTAAGGT	[945]
030_EF1a_Chaetomorpha_coliformis	--AAGGAAT	CAAGCGTGGT	TTGCTGCGCT	CGGACTCCAA	G-----	AAT--GACC	CC--GCCAA	GGGCTGC---	-----	T AGTACTTCA	CGCCACAGGT	[795]
031_EF1a_Arabidopsis_thaliana	--AAGGATCT	TAAGAGAGGA	TACGTTGGCT	CTAAGTCCAA	G-----	GAT--GATC	CA--GCTAA	GGGTGCC---	-----	G CCAACTTAC	CTCCACAGGT	[996]
032_EF1a_Oryza_sativa	--AAGGATCT	GAGCGTGGG	CCAAGTGGCT	CCAAGTCCAA	G-----	GAT--GACC	CT--GCCAA	GGAGGCT---	-----	G CCAGCTTAC	CTCCACAGGT	[996]
033_EF1a_Cyanidioschyzon_merolae	--AAGGAAT	GAGCGCGGC	TTTGTATGCG	GAGACAGCAA	G-----	AAC--GACC	CA--CCCAA	GGCCGCG---	-----	G AGGACTTCA	GGCACAGGT	[1008]
034_EF1a_Porphyrha_jezoensis	--AAGGACCT	GAGCGCGGC	TACGTTGGCG	GTGACTCGAA	G-----	AAC--GACC	Cg--CCGAA	GGGGTGC---	-----	G CTTCGTTCA	CGCCACAGGT	[1008]
035_EF1a_Plasmodium_falciparum	--AAGAAAT	CAAGAGAGGT	TACGTTGGCT	CAGATACTAA	A-----	AAT--GACC	CA--GCCAA	AGGATGC---	-----	T CCAACTTAC	TGCTCAAGTT	[990]
036_EF1a_Cryptosporidium_parvum	--AAGGATAT	CAAGAGAGGT	TTGTTGGCTT	CTGATGCCAA	G-----	AAT--GACC	CA--GCTAA	GGGCTGT---	-----	G AAGACTTAC	TGCTCAAGTT	[984]
037_EF1a_Tetrahymena_pyriformis	--AAGGATAT	CAGAAGAGGT	AATGTCGCTT	CGGATGCCAA	G-----	AAC--GACC	CC--GCCAA	AGAGGCT---	-----	G CCAGCTTCA	CTCTCAAGTC	[999]
038_EF1a_Stentor_coeruleus	--AAGGATAT	CAAGAGAGGA	TACGTCGCTT	CTGATGCCAA	A-----	GCT--GACC	CT--GCTAA	AGAAATGT---	-----	G AAAACCTAC	TGCTCAAGTC	[951]
039_EF1a_Phytophthora_parasitica	--AAGGAGCT	CGCTCGTGGT	CGGACTCGCT	CGGACTCCAA	G-----	AAC--GACC	Cg--GCCAA	GGGACCT---	-----	C AGGACTTAC	CGCCACAGGT	[996]
040_EF1a_Mallomonas_rasilis	--AAGGATAT	TAAGAGAGGA	TCNGTCGCTT	CCAAGTCTGC	T-----	GAG--GACC	CT--GCCAA	GGGATGT---	-----	G ATACTTTCAC	TGCACAAGTT	[957]
041_EF1a_Phaeodactylum_tricornutum	--AAGGATCT	CGCGCTGGC	TTTGTGCTGT	CTGACTCGAA	A-----	GCT--AGCC	CC--GCTTC	GGGTGTT---	-----	T CTTCTTTCGA	GGCTCAGGTT	[996]
042_EF1a_Thalassiosira_pseudonana	--CAAGATCT	TCATCGAGGT	GATGTTGCTA	GTGAGGCAGG	A-----	AAG--AAGC	AA--GCCAC	AGCTGCC---	-----	A AGTCTTTGA	TGCCACAGCT	[996]
043_EF1a_Malawimonas_jakobiformis	--AAGGATTT	CGCGCTGGC	TTGTTTGGCG	GTGACACCAA	G-----	AAC--GACC	CC--CCGAT	GGAGGCT---	-----	G GCGAGTTAC	TGCTCAGGTT	[927]
044_EF1a_Malawimonas_californiana	--AAGGATAT	CGCGCTGGC	AACGTTGCGC	GTGACTCCAA	G-----	AAC--GACC	CC--CCCAA	GGAGTCT---	-----	G CGGACTTAC	CGCCACAGGT	[996]
045_EF1a_Trimastix_pyriformis	--AAGGATCT	GAGCGTGGC	TTGCTGCTGT	CGGACTCCCG	T-----	GTT--GACC	CC--CCTCA	GGAAACT---	-----	G AGAAGTTCT	CGCCACAGGT	[954]
046_EF1a_Dinenympha_exilis	--AAGATTT	GAAACGTGGA	TTGTTGCTGT	GAGATGCCAA	G-----	AAT--GACC	CA--GCCAA	AGAAAGCT---	-----	G ACACATTCT	TGCACAAGTT	[936]
047_EF1a_Monocercomonoides_sp.	--AAGGATAT	CGCTCGTGGG	AATGTTGCTG	GAGACTCCAA	A-----	CAA--GATC	CA--CCCAT	GGAAAGCT---	-----	G AATCATTCGT	TGCACAAGTT	[996]
048_EF1a_Streblospio_trix	--AAGGATAT	TAAGAGAGGA	TATGATATGT	GAGATTCTAA	A-----	TAA--GATC	CA--CCTCA	AGAAAGCT---	-----	G AATCATCTA	AGCACAGGT	[945]
049_EF1a_Histomonas_meleagridis	--TCTGACAT	CAAGCGTGGT	TACGTCGCTG	GTGAAGCTAA	C-----	CGT--GATC	CA--CCAAT	GCAATGC---	-----	A TCAACTTAC	AGCACAAATG	[996]
050_EF1a_Trichomonas_vaginalis	--GCTGATGT	CAAGCGTGGC	TACGTCGTTG	GTGATACAAA	G-----	CGT--GACC	CA--CCAGT	GCAATGC---	-----	G CTTCTTTCAC	AGCTCAGATG	[996]
051_EF1a_Hexamita_inflata	--AAGGACAT	CAAGAGAGGT	TACGTCGCTG	GTGACTCCAA	C-----	AAT--GAGC	CT--CCCAA	GGGCTGC---	-----	G AATACTTCA	GGCTAACGTC	[948]
052_EF1a_Spiroplasma_vortens	--AAGGACAT	CAAGAGAGGT	TACGTCGCTG	GTGACTCCAA	G-----	GAGTCCGACC	CA--CCAGG	CAGAGTG---	-----	A AGTCTTTCG	GGCCACAGGT	[951]
053_EF1a_Amaluella_incarcerata	--AAGGATAT	CGCGCTGGT	AACGTTGCTT	CTGACACCAA	G-----	AAC--GACC	CT--GCCAA	GGCTTGC---	-----	G AATCTTTCGT	CGCTCAGGTC	[945]
054_EF1a_Seculamonas_ecuadoriensis	--AAGGACAT	CGCGCTGGC	TACGTCGCTG	CGGACTCCAA	G-----	AAC--GACC	Cg--GCCAA	GGAGGCT---	-----	G CCTCGTTCA	CGCCACAGGT	[996]
055_EF1a_Reclinomonas_americana	--AAGGATAT	CGCGCTGGC	TACGTCGCTG	CGGACTCCAA	G-----	AAC--GACC	CA--GCCAA	GGAGGCT---	-----	A CTTCTTTCAC	TGCCACAGGT	[996]
056_EF1a_Histiona_roides	--AAGGATAT	CGCGCTGGC	TACGTTGGCT	GTGACACCAA	G-----	AAC--GACC	CC--GCCAA	GGAGGCT---	-----	G AGTCTTTCAC	TGCCACAGGT	[996]
057_EF1a_Jakoba_libera	--AAGGATAT	CGCGCTGGC	AACGTCGCTG	CGGACTCCAA	G-----	AAC--GACC	CC--GCCAA	GGAGTCC---	-----	A CCAGCTTCA	CGCCACAGGT	[996]
058_EF1a_Stachyamoeba_lipophora	--AAGGATGT	CGCGCTGGC	TTGTTGCTGT	GTGAGACCAA	G-----	AAC--GACC	CC--CCTAA	GGAGTCT---	-----	G CTTCTTTCAC	TGCCACAGGT	[996]
059_EF1a_Naegleria_gruberi	--ACTGATAT	TAAGAGAGGT	TATGTTGCTT	CTGATGCTAA	G-----	AAT--GATC	CA--TCCAG	AGAAAGCT---	-----	G TTTCTTTCAC	TGCTCAAGTC	[948]
060_EF1a_Naegleria_andersoni	--ACTGATAT	CAAGAGAGGT	TATGTTGCTT	CTGATGCTAA	G-----	AGT--GATC	CT--TCAAG	AGAAAGCT---	-----	G TTTCTTTCAC	TGCTCAAGTT	[966]
061_EF1a_Acrasis_rosea	--AAGGATAT	CGCGAGAGGA	AACGTTTGGT	CAAAATCTAA	G-----	AAC--GACC	CA--CCGAG	AGGTTGC---	-----	G AGAACTTTC	TGCCACAGGT	[945]
062_EF1a_Sawyeria_marylandensis	--AAGGATAT	TAAGAGAGGT	AACGTTTGGC	GTGACATTAA	A-----	AAT--GATC	CA--CCCAA	AGAAAGCT---	-----	G CAAGCTTGT	TGCCACAGGT	[996]
063_EF1a_Bodo_saltans	--AAGGATAT	CGCTCGTGGC	TTGCTGCTGC	GTTCGCGCAA	G-----	GTC--GACC	CC--CCGAA	GGAGTGC---	-----	G AGAGCTTAC	CGCTCAGGT	[936]
064_EF1a_Rhynchobodo_sp.	--AAGGAGCT	CAAGCGTGGG	TTGCTGCTGC	GTACACCCAA	G-----	CAG--GACC	CA--CCCAA	GGGATGC---	-----	G ACTTCTTAC	CGCCACAGGT	[930]
065_EF1a_Rhynchomonas_nasuta	--AAGGACAT	CGCGCGCGG	TTGCTGCTGC	CCAAGTCCAA	G-----	AAC--GACC	CC--GCCAA	GGAGGCC---	-----	G AGTCTTCTC	CGCCACAGGT	[993]
066_EF1a_Dimastigella_trypaniformis	--AAGGACAT	CGCTCGTGGC	TTGCTGCTGC	CGAAGTCCAA	G-----	AAC--GACC	CC--GCCAA	GGAGGCC---	-----	G AGAAGTCTC	GGCCACAGGT	[951]

	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	
067_EF1a_Perkinsiella_amoebae	--AAGGACAT	CGGCCGTGGT	TTTGTCTGCT	GGAACAGCAA	G-----	CAG---GACC	CC---GCCAA	GGAGTGC---	-----G	AGCACTTCAT	CGCG-----	[915]
068_EF1a_Leishmania_major	--AAGGACAT	CGGCCGTGGT	AACGTGTGCG	GCAACTCGAA	G-----	AAC---GACC	CG---CCGAA	GGAGGCC---	-----G	CCGACTTCAC	GGCGCAGGTG	[996]
069_EF1a_Leishmania_infantum	--AAGGACAT	CGGCCGTGGC	AACGTGTGCG	GCAACTCGAA	G-----	AAC---GACC	CG---CCGAA	GGAGGCC---	-----G	CCGACTTCAC	GGCGCAGGTG	[996]
070_EF1a_Leishmania_braziliensis	--AAGGACAT	CGGCCGTGGT	AACGTGTGCG	GCAACTCGAA	G-----	AAC---GACC	CG---CCGAA	GGAGGCC---	-----G	CCGACTTCAC	GGCGCAGGTG	[996]
071_EF1a_Leishmania_tarentolae	--AAGGACAT	CGGCCGTGGG	AACGTGTGCG	GTAACCTCGAA	G-----	AAC---GACC	CG---CCGAA	GGAGGCC---	-----G	CCGACTTCAC	GGCGCAGGTG	[996]
072_EF1a "L. tarentolae" FJ807257	--AAGGACAT	CGGCCGTGGG	AACGTGTGCG	GTAACCTCGAA	G-----	AAC---GACC	CG---CCGAA	GGAGGCC---	-----G	CTGACTTCAC	GGCG-----	[915]
073_EF1a_Leptomonas_costaricensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[594]
074_EF1a_Leptomonas_bifurcata	--AAGGACAT	CGGCCGTGGT	AACGTGTGCG	GCAACTCGAA	G-----	AAC---GACC	CG---CCGAA	GGAGGCC---	-----G	CCGACTTCAC	CGCTCAGGTG	[756]
075_EF1a_Leptomonas_podlipaevi	--AAGGACAT	CGGCCGTGGC	AACGTGTGCG	GTAACCTCGAA	G-----	AAC---GACC	CG---CCGAA	GGAGGCC---	-----G	CCGACTTCAC	CGCCACAGGTG	[744]
076_EF1a_Herpetomonas_muscarum	--AAGGACAT	CGGCCGTGGT	AACGTGTGCT	CCAACCTCAA	G-----	AAC---GACC	CC---GCCCG	CGAGGCC---	-----G	AGGACTTCAC	CGCCACAGGTG	[909]
077_EF1a_Herpetomonas_pestoi	--AAGGATAT	CGGCCGTGGT	AACGTGTGCT	CCAACGCCAA	G-----	AAC---GACC	CC---GCTCG	CGAGGCT---	-----G	AGGACTTCAC	CGCC-----	[771]
078_EF1a_Blastocystidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[766]
079_EF1a_Trypanosoma_brucei	--AAGGACAT	CGGCCGTGGC	AACGTCTGCG	GTAACACCAA	G-----	AAC---GACC	CC---CCAAA	GGAGGCC---	-----G	CCGACTTCAC	GGCACAGGTG	[996]
080_EF1a_T. brucei_equiperdum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[774]
081_EF1a_Trypanosoma_aviium	--AAGGACAT	CGGTCGTGGT	AATGCTCTGT	GTAACCTCGAA	G-----	AAC---GACC	CC---CCGAA	GGAGGCC---	-----G	CCGACTTCAC	GGCGCAGGTG	[942]
082_EF1a_Trypanosoma_cruzi	--AAGGACAT	CGGCCGTGGC	AACGTGTGCG	GCAACTCGAA	G-----	AAC---GACC	CC---CCAAA	GGAGGCC---	-----G	CCGACTTCAC	GGCGCAGGTG	[996]
083_EF1a_Trypanosoma_rangeli	--AAGGACAT	CGGTCGTGGC	AACGTGTGCG	GCAACTCGAA	G-----	AAC---GACC	CT---CCGAA	GGAGGCC---	-----G	CCGACTTCAC	GGCGCAGGTG	[996]
084_EF1a_Ploetitia_costatum	--AAGGATAT	TCGTCGTGGT	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCC---	-----G	CCGACTTCAC	CGCTCAGGTG	[993]
085_EF1a_Ploetitia_oblonga_cDNA	--AAGGATAT	TCGTCGTGGG	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCC---	-----G	CCGACTTCAC	CGCTCAGGTG	[993]
086_EF1a_Ploetitia_oblonga_gDNA	--AAGGATCT	CGGCCGTGGG	TTGTTGTGCT	CCAACGCCAA	G-----	CAG---GACC	CC---GCCAA	GGAGGCT---	-----G	AGGACTTCAC	CGCTCAGGTG	[993]
087_EF1a_Entosiphon_sp.	--AAGGAAAT	CGGCCGTGGC	TTGTTGTGCT	GGAACGCCAA	G-----	AAC---GACC	CG---GCCAA	GGAGGCT---	-----G	AGGACTTCAC	TGCCAACGTT	[993]
088_EF1a_Entosiphon_sulcatum	--AAGGAAAT	TCGCCGTGGC	TTGTTGTGCT	GGAACGCCAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	AGGACTTCAC	TGCT-----	[915]
089_EF1a_Peranema_tricophorum	--AAGGATAT	CAGACGTGGA	TTGTTGCT--	-----	-----	-----	-----	-----	-----	-----	-----	[938]
090_EF1a_P. tricophorum_FJ807243	--AAGGATAT	CAGACGTGGA	TTGTTGCTCT	CAAACCTCAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCTCAGGTT	[987]
091_EF1a_Colacium_elongatum	--AAGGATAT	CGGACGTGGC	TACGTGGCTT	CTAATGCCAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCCACAGGTG	[993]
092_EF1a_Colacium_mucronatum	--AAGGACAT	CGGACGTGGC	TACGTGGCTT	CTAATGCCAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCCACAGGTG	[939]
093_EF1a_Colacium_sp.	--AAGGATAT	CGGTCGTGGC	TACGTTTGCT	CCAACGCCAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCCACAGGTG	[993]
094_EF1a_Colacium_vesiculosum	--AAGGACAT	CGGACGTGGC	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCCACAGGTG	[951]
095_EF1a_Cryptoglena_pigra	--AAGGATAT	TCGTCGTGGC	TATGTTGTCT	CCAATTCGAA	G-----	AAC---GACC	CG---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCCACAGGTG	[993]
096_EF1a_Cryptoglena_skujai	--AAGGATAT	TCGACGTGGC	TTGTTGTGCT	CCAATTCGAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCCACAGGTG	[993]
097_EF1a_Discoplastis_adunca	--AAGGATAT	TCGACGTGGC	TACGTGGCTT	CCAACGCCAA	G-----	AAT---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCTCAGGTT	[993]
098_EF1a_Euglena_clara	--AAGGATAT	TCGCCGTGGG	TTGTTGTGCT	CTAATTCGAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCCAACGTT	[873]
099_EF1a_Euglena_deses	--AAGGATAT	CGGACGTGGC	TATGTTGTCT	CCAATTCGAA	G-----	AAT---GACC	CT---GCCAA	AGAGGCT---	-----G	GTGACTTCAC	TGCTCAGGTT	[993]
100_EF1a_Euglena_exilis	--AAGGATAT	CGGTCGTGGG	TTGTTGTGCT	CCAATTCGAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACGTT	[873]
101_EF1a_Euglena_geniculata	--AAGGATAT	TCGTCGTGGG	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCTCAGGTT	[993]
102_EF1a_Euglena_gracilis	--AAGGATAT	TCGTCGTGGG	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACGTT	[993]
103_EF1a_E. gracilis_X16890	--AAGGATAT	TCGTCGTGGG	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACGTT	[996]
104_EF1a_E. gracilis_FJ807238	--AAGGATAT	TCGTCGTGGG	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACGTT	[996]
105_EF1a_Euglena_granulata	--AAGGATAT	TCGACGTGGA	TATGTTGTCT	CCAACGCCAA	G-----	AAT---GACC	CC---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCCAACGTT	[993]
106_EF1a_Euglena_hiemalis	--AAGGATAT	CGGTCGTGGT	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCCAACGTT	[993]
107_EF1a_Euglena_longa	--AAGGATAT	CGGTCGTGGT	TATGTTGTCT	CCAATTCGAA	G-----	AAC---GACC	CA---GCCAA	GGAGGCT---	-----G	CCGACTTCAC	TGCCAACGTT	[510]
108_EF1a_Euglena_stellata	--AAGGATAT	TCGACGAGGT	TTGTTGTGCT	CCAATTCGAA	G-----	AAT---GACC	CA---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCTCAGGTT	[873]
109_EF1a_Euglena_viridis	--AAGGATAT	CGGACGTGGA	TTGTTGTGCT	CCAACGCCAA	G-----	AAT---GACC	CA---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACGTT	[951]
110_EF1a_Eutreptia_pertyi	--AAGGATAT	TCGACGTGGA	TATGTTGTCT	CCAACGCCAA	G-----	AAT---GACC	CC---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCCAACGTT	[993]
111_EF1a_Eutreptiella_pomquetensis	--AAGGATAT	TCGTCGTGGG	TATGTTGTCT	CCAACGCCAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	CCGACTTCAC	CGCTCAGGTT	[603]
112_EF1a_Lepocinlis_acus_var._major	--AAGGATAT	CGGCCGTGGC	TACGTGTGCT	CTAACTCCAA	G-----	AAT---GACC	CG---GCCAA	GGAGGCT---	-----G	CGGACTTCAC	TGCCAACGTT	[993]
113_EF1a_Lepocinlis_ovum	--AAGGATAT	CGGCCGTGGC	TATGTTGTCT	CTAACTCCAA	G-----	AAT---GACC	CT---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCTCAGGTT	[873]
114_EF1a_Lepocinlis_spyrogyroides_clone_38	--AAGGATAT	TCGCCGTGGG	TTTGTGTGCT	CCAATTCGTA	G-----	AAC---GACC	CT---GCCAA	AGAGGCT---	-----G	CTGACTTCAC	CGCTCAGGTT	[306]
115_EF1a_L. spyrogyroides_clones_1_&_4	--AAGGATAT	TCGCCGTGGG	TACGTGTGCT	CCAATTCGTA	G-----	AAT---GACC	CT---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCCAACGTT	[993]
116_EF1a_Lepocinlis_steinii	--AAGGATAT	CGGCCGTGGC	TTTGTGTGCT	CCAACGCCTA	G-----	AAT---GACC	CT---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	AGCCACAGGTG	[873]
117_EF1a_Lepocinlis_triaperis	--AAGGATAT	CGGTCGTGGT	TACGTGTGCT	CCAATTCGTA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	CGGACTTCAC	TGCCAACAGGTG	[993]
118_EF1a_Monomorphina_aenigmatica	--AAGGATAT	CGGCCGTGGG	TACGTGTGCT	CCAATTCGTA	G-----	AAT---GACC	CA---GCCAA	AGAGGCT---	-----G	CTGACTTCAC	CGCCACAGGTG	[993]
119_EF1a_Monomorphina_pyrum	--AAGGATAT	TCGCCGTGGC	TACGTGTGCT	CCAATTCGTA	G-----	AAT---GACC	CT---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCCAACGTT	[993]
120_EF1a_Monomorphina_sp.	--AAGGACAT	CGGTCGTGGT	TACGTGTGCT	CCAACGCCAA	G-----	AAT---GACC	CT---GCCAA	GGAGGCT---	-----G	CGGACTTCAC	CGCTCAGGTT	[993]
121_EF1a_Phacus_orbicularis	--AAGGATAT	TCGTCGTGGA	TTTGTGTGCT	CCAATTCGTA	G-----	AAT---GACC	CA---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	TGCCAACGTT	[951]
122_EF1a_Phacus_similis	--AAGGATAT	CGGCCGTGGC	TTGTTGTGCT	CCAACGCCAA	G-----	AAC---GACC	CG---GCCAA	GGAGGCT---	-----G	CTGACTTCAC	CGCCACAGGTG	[993]
123_EF1a_Strombomonas_acuminata	--AAGGACAT	TCGTCGAGGA	TATGTTGTCT	CCAATTCGTA	G-----	AAT---GACC	CC---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACAGGTG	[996]
124_EF1a_Strombomonas_borysteniensis	--AAGGATAT	CGGCCGTGGC	TACGTGTGCT	CCAATTCGTA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACAGGTG	[873]
125_EF1a_Strombomonas_napiformis	--AAGGATAT	CGGTCGTGGA	TATGTTGTCT	CCAATTCGTA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	CGCCACAGGTG	[993]
126_EF1a_Strombomonas_ovalis	--AAGGATAT	TCGCCGTGGC	TACGTGTGCT	CCAACGCCAA	G-----	AAC---GACC	CC---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACAGGTG	[993]
127_EF1a_Strombomonas_triquetra	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[891]
128_EF1a_Trachelomonas_conradi	--AAGGATAT	TCGACGTGGG	TATGTTGTGCT	CCAACGCCAA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GGGACTTCAC	TGCCAACAGGTG	[993]
129_EF1a_Trachelomonas_echinata	--AAGGATAT	TCGACGAGGC	TATGTTGTGCT	CCAATTCGTA	G-----	AAT---GACC	CA---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCTCAGGTT	[993]
130_EF1a_Trachelomonas_lefevrei	--AAGGATAT	CGGACGTGGA	TTGTTGTGCT	CCAATTCGTA	G-----	AAC---GACC	CT---GCCAA	GGAGGCT---	-----G	GGGACTTCAC	TGCCAACAGGTG	[993]
131_EF1a_Trachelomonas_planchetonia	--AAGGATAT	CGGACGTGGA	TTGTTGTGCT	CCAATTCGTA	G-----	AAT---GACC	CC---GCCAA	GGAGGCT---	-----G	GGGACTTCAC	TGCCAACAGGTG	[873]
132_EF1a_Trachelomonas_similis	--AAGGATAT	TCGACGAGGT	TTTGTGTGCT	CCAATTCGTA	G-----	AAT---GACC	CT---GCCAA	GGAGGCT---	-----G	GTGACTTCAC	TGCCAACAGGTG	[873]

	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	
133_EFL1a_Trachelomonas_volvocina	--AAGGACAT	TCGTCTGTGG	TTTGTCTGCT	CAAAACGCAA	G-----	AAT---GACC	CT---GCGAA	GGAGGCT---	-----G	GTGATTTCAC	TGCACAGGTC	[993]
134_EFLenvironmental_sample_EU810232	GCAACAAGCC	CAGGACCGGA	GACATCATGA	TCTTCAAGTC	C-----	GAT---ACCT	CC---CTCAA	GCCTTGC---	-----A	AGAGCTTCGT	CGCCAGGTC	[1086]
135_EFLCreolimax_fragrantissima	AGAACATGCC	CCGCGCTGGT	GATGTCATGA	TCTTGAAGTC	T-----	GAC---ACCA	CT---CTTGG	ACGATGC---	-----A	-----	-----	[994]
136_EFLSphaeroforma_arctica	AGAACATGCC	TCGCGCTGGA	GATGTCATGA	TTCTAAAGTC	C-----	GAT---AACA	CT---CTAGG	TCGATGC---	-----A	AGAACTTCAC	AGCACAAGTA	[1086]
137_EFLMonosiga_brevicollis	ACAACATGCC	CCGTGTTGGT	GATGTCATGA	TCTTGAAGTC	T-----	GAT---AGCT	CC---ATTGG	CCGCTGC---	-----A	AGTCTTTCAC	CGTCCAGGTC	[1086]
138_EFLXenopus_laevis	AGAATATGCC	CGCGACTGGT	GACGTCATGA	TCTTAAATATC	G-----	GAC---CAGA	GC---TTGGG	CCGCTGC---	-----A	AGAACTTCAC	CGTCCAGGTC	[1086]
139_EFLSpizellomyces_punctatus	ACAACATGCC	CCGTGCTGGT	GACATCATGA	TCTACAAGGA	G-----	GAC---AAGT	CC---TTGAA	GCCCGTC---	-----A	AGGCTTTCAC	TGCCAAGTC	[1092]
140_EFLGaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
141_EFLAllomyces_macrogynus	ACAACATGCC	CCGCGTGGT	GACGTCATGA	TCTACAAGGA	G-----	GAC---TCGA	CG---CTCAA	GGCCGTC---	-----A	AGTCGTTTAC	CGCCAGGTC	[1092]
142_EFLBlastocladiella_emersonii	ACAACATGCC	CCGCGTGGT	GACGTCATGA	TCTACAAGGA	G-----	GAC---GCCA	CC---CTCAA	GGCCGTC---	-----A	AGTCGTTTAC	CGCCAGGTC	[1092]
143_EFLCatenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
144_EFLCoelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
145_EFLEntomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
146_EFLBasidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
147_EFLConidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
148_EFLTetraselmis_tetratele	ACAACATGCC	CCGCTCCGGT	GACGTGATGA	TCCACAAGTC	C-----	GAC---ACCA	CC---CTAGG	CCATGTG---	-----G	AGGAAGTTTAC	TGCCAAGATC	[1035]
149_EFLMesostigma_viride	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[879]
150_EFLOstreococcus_tauri	TCAACATGCC	CGCGACCGGT	GACGTCATGA	TCTACAAGAA	G-----	GAC---ACCT	CC---CTCGC	CCCCTGC---	-----A	AGAACTTCAC	CGCGCAAGTC	[1095]
151_EFLNephrolepis_olivacea	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[957]
152_EFLMicromonas_pusilla	TCAACATGCC	CCGCTCCGGT	GACGTCATGA	TCTACAAGAA	G-----	GAC---ACCA	CC---CTCAA	GGCGTGC---	-----A	AGAACTTCAA	CGCCAGATC	[1044]
153_EFLParachlorella_kesslerii	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[942]
154_EFLHelicosporidium_sp.	ACAACATGCC	CCGCGACCGG	GACGTCATGA	TCATCAAGTC	T-----	GAC---GCCA	CC---CTCGC	CCCCGTC---	-----A	AGGAAGTTTAC	CGCCAGATC	[1107]
155_EFLGonium_pectorale	GCAACATGCC	CCGCGACCGGT	GACGTGATGA	TCTTCAAGTC	G-----	GAC---AGCA	CC---CTGAA	GCTGGTG---	-----C	AGAGCTTTCAC	CGCCACCATC	[1089]
156_EFLChlamydomonas_reinhardtii	GCAACATGCC	CCGCGACCGGT	GACGTGATGA	TCTTGAAGTC	G-----	GAC---CAGA	CC---CTGAA	GATCGTC---	-----A	AGGAAGTTTAC	CGCCAGATT	[1089]
157_EFLScenedesmus_obliquus	GTAACATGCC	CCGCGACCGG	GACGTGATGA	TCTTGAAGAG	C-----	GAC---CAGA	CC---CTGAA	GCCCTGC---	-----A	AGGAAGTTTAC	CGCCAGATC	[1089]
158_EFL"Oryza_sativa"	GCAACATGCC	CCGCGACCGGT	GACGTGATGA	TCTTGAAGAG	C-----	GAC---CAGA	CC---CTGAA	GCCATGC---	-----C	AGAACTTTCAC	TGCACAGATC	[1089]
159_EFLChlorococcum_sp.	GCAACATGCC	CCGTGCTGGT	GATGTTATGA	TCCACAAGAA	C-----	GAT---GGCA	CC---CTCAA	GCACGTT---	-----G	CCAAGTTTAC	TGCCAGATT	[1038]
160_EFLOchlochaete_hystrix	ACATGATGCC	CCGCGACCGG	GACGTGATGA	TCTTCAAGTC	C-----	GAC---ACCA	CC---CTGAA	GCAGGTG---	-----A	AGGAAGTTTAC	CGCCAGATC	[1047]
161_EFLUlva_intestinalis	GCAACATGCC	CCGCGAGGGT	GATGTCATGA	TCATGAAGTC	C-----	GAT---CCTT	CT---CTGAA	GCACGTC---	-----G	AGGAAGTTTCT	TGCTCAGATC	[1038]
162_EFLUrospora_sp.	GTAACATGCC	CCGCGCGGGG	GACGTGATGA	TCATGAAGAG	C-----	GAC---GCCA	GC---CTCAA	GCACGTG---	-----G	AGGAAGTTTAC	CGCCAGATC	[891]
163_EFLGymnochlorella_stellata_ppc	ATAACCTCTC	TGATGCTGGA	GATGTTGATGA	TATTGAGTGG	A-----	GAT---AAGA	GT---TTACA	GTCACAT---	-----G	CGGAGTTTAC	AGCACAAGTC	[1125]
164_EFLLotharella_amoebiformis_ppc	ACAATCTCTC	TAGATCCGGT	GATGTTGATGA	TCTTGAAGTC	G-----	GAC---AAGA	GC---CTCAA	GTCYACC---	-----G	AGGAAGTTTAC	CGCCAGGTC	[870]
165_EFLChlorarachnion_reptans_ppc	ACAACCCCTC	CCGCTCTGGC	GATGTTGATGA	TCTTGAAGTC	G-----	GAT---GATT	CC---CTCCA	TGCTACT---	-----G	GTGAATTTAC	AGCGAGGTT	[897]
166_EFLChondrus_crispus	TAAACATGCC	TCGCGTCCGT	GATGTTGATGA	TTCTACCAAA	G-----	GAT---GACT	CC---CTCAA	CGGTTGC---	-----G	CCAGTTTTCAC	GGCCAGGTT	[1086]
167_EFLGracilaria_changii	TGAACATGCC	TACGTCCTGG	GATGTTGATGA	TTCTCCAGAA	G-----	GAT---ACTA	CC---TTGAA	GCACACA---	-----G	CAAACTTTCAC	TGCCAGGTT	[1086]
168_EFLCryptomonas_ovata	ACAACATGCC	CCGCTCTGGC	GACGTGATGG	TGTACGCCAA	G-----	GAC---GGCA	CC---CTCAA	GGGACCC---	-----A	AGTCCTTTCAC	TGCTCAGATC	[1044]
169_EFLGuillardia_theta	ACAACATGCC	CCGCTCTGGC	GACGTGATGG	TGTACAAGAA	G-----	GAT---GGTA	CC---CTCAA	GGGAACC---	-----A	AGTCCTTTCAC	TGCACAGATC	[1089]
170_EFLRhodomonas_salina	ACAACATGCC	CCGTGCGGGT	GACGTTATGG	TGTACGCCAA	G-----	GAC---GGCA	CC---CTCAA	GGGACCC---	-----A	AGTCCTTTCAC	TGCCAGATC	[1086]
171_EFLPavlova_lutheri	AGAACATGCC	CGCGACCGG	GACGTGATGG	TGTACAAGAA	G-----	GAC---GGCA	CG---CTCAA	GGCGGCC---	-----G	AGAACTTTCAC	CGCCAGATC	[1086]
172_EFLChrysochromulina_sp.	ACAACATGCC	CGCGCTTGGC	GACGTGATGA	TCTACAAGAA	G-----	GAC---TCGT	CG---CTCGA	CCGCTGC---	-----G	GCCAGTTTAC	CGCGAGATC	[906]
173_EFLIsochrysis_galbana	TCAACATGCC	CCGCGTCCGG	GACGTGATGA	TCTACAAGAA	G-----	GAC---AAGT	CG---CTGGA	CCAGGTG---	-----G	GCCAGTTTCA	CGCGAGGTC	[1089]
174_EFLEmiliana_huxleyi	AGAACATGCC	CCGCGTCCGG	GACGTGATGA	TCTACAAGAA	G-----	GAC---AAGA	CG---CTGAC	GCAGTGC---	-----G	CGGAGTTTCA	CGCGAGATC	[1095]
175_EFLAchnanthes_kuwaitensis	ACGAGAAAGT	CCAGCCTGGC	GATATTTATTT	ACATCCAGAA	G-----	GAG---GGCG	AT---CTTAA	GCCCATC---	-----A	AGTCGTTTAC	TGCCATGGTC	[996]
176_EFLAsterionella_glacialis	ACGAGAAAGT	CTCTCCCGGA	GACATTATCT	ACCTCCAGAA	G-----	GAG---GGAG	AG---TGCTT	GCCTGTC---	-----A	AGCTTTTTCAG	AGCCATGGTT	[996]
177_EFLDetonula_confervacea	ACGAGAAAGT	CTCCCTTGGA	GATATCATCT	ACGTCAGAG	T-----	GAA---GGAG	AC---CTCAA	GTCATCC---	-----A	AGTCCTTTCAC	TGCTGCTGTT	[996]
178_EFLThalassiosira_pseudonana	ACGAGAAAGT	CCAGCCTGGA	GATATCATCT	ACGTCAGAG	G-----	GAG---GGAG	AT---CTCAA	ACCCATC---	-----A	AGTCCTTTCAC	TGCCATGGTT	[1059]
179_EFLPerkinsus_marinus	AGAATATGCC	TCGTTCTGGA	GATGTTATGG	TCTACAAGAA	A-----	GAT---ACCA	CT---CTTAA	ACATGTG---	-----C	AATCTTTCAC	TGCTCAGGTC	[1104]
180_EFLOxyrrhis_marina	AGAACATGCC	CCGCGCTGGT	GATGTTGATGA	TTTACAAGAA	G-----	GAC---ACCA	CG---CTCTC	TGCGGCT---	-----G	GCAGCTTTCAC	CGCAGAGTT	[1098]
181_EFLKarlodinium_micrum	ACAACATGCC	ACGCTCTGGC	GATGTTGATGG	TGTACAAGAA	G-----	GAT---ACGA	CC---CTCGG	ACAAACA---	-----A	AGGAAGTTTCA	TGCACAGATT	[1092]
182_EFLAmphidinium_carterae	ACAACATGCC	AAGGTCAAGT	GATGTTGATGG	TGTACAAGAA	G-----	GAC---TCTA	CC---TTGGG	TCAGACC---	-----A	AGGCCTTTCAG	TGCACAGATC	[1092]
183_EFLHeterocapsa_triquetra	ACAACATGCC	CGGATCGGGG	GACGTGATGG	TGTACAAGAA	G-----	GAC---ACGA	CG---CTGGG	CCAGACT---	-----C	GCAGGTTTCA	TGCCAGATC	[1092]
184_EFLPlanoglabratella_operculis	ATAATATGCC	GATGTTTGGT	GATGTTAATGG	CCATTGACGA	T-----	GAT---CCAG	AGGCCACCTAA	ACAAGCC---	-----T	CAGCTTTTAC	GGCTTTGGTA	[924]
185_EFLReticulomyxa_filosa	AAAATATGCC	ACACATGATG	GCATTGATGA	CCCTGCAAGT	GAT---	CCCTA	ACCCACCAAA	ACAAGCA---	-----A	CCAAATTTTCA	GGCTTTGGTA	[999]
186_EFLLotharella_amoebiformis	ACCGTCTCCC	CAAGGTTGGA	GAGATCATGG	CCCTTCAGGA	T-----	-----GATA	CCACTATCGG	AAAAACT---	-----A	AGTCCTTTCAC	TGCCAGGTA	[1167]
187_EFLGymnochlorella_stellata	GAGTCTTCCC	TAAGGTTGGA	GAAGTTATGG	TCTTTCAGTC	T-----	GAG---GATT	TCGGCAGCCC	TCTTCTCTC	GGAAAAACCA	AGAAGTTTCTG	TGTATCTGTC	[1140]
188_EFLChlorarachnion_reptans	AGGTCTTCCC	AAAGTTGGA	GAAGTTATGA	CTTTGGCCGA	C-----	GAT---GACT	TTGGAAGCCC	TCCCAATTC	GGAAAAACAA	AGATTTTCTGTC	GGTTTCTGTC	[1071]
189_EFLDiplonema_ambulator	ACAACATGCC	CAGGACCGGT	GACGTGATGG	TGTACAAGAA	G-----	GAC---GGGT	CC---CTCAA	GGAGGCC---	-----G	GTACCTTTCAC	CGCCAGGTC	[690]
190_EFLDiplonema_papillatum	ACAACATGCC	CCGAGCCGGA	GACGTGATGG	TGTACAAGAA	G-----	GAC---AGCT	CG---CTGAA	GGAGGCT---	-----G	GCAGGTTTCA	GGCCAGGTC	[1062]
191_EFLRhynchopus_euleeides	ACAACATGCC	CCGAGCCGGT	GACGTGATGG	TGTACAAGAA	G-----	GAC---GCCA	CC---CTGAA	GGAGGCT---	-----G	GCAGCTTTCAC	GGCCAGGTC	[705]
192_EFLPetalomonas_cantuscygni	AGAACATGCC	CGGACTGGC	GACGTGATGG	TCTTCAACAA	G-----	GAT---CCCT	CC---CTCAA	GTCCGCG---	-----G	GCCAGTTTAC	C-----	[924]
193_EFLTrypanoplasma_borrelia	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[1017]
194_EFLNeobodo_salienis	CCAACATGCC	GCACGACGGT	GACATCATGG	TCAAGAAGAG	C-----	GAC---ACCG	AC---TTCAA	GGCGTGC---	-----G	AGTCCTTCTG	CGTCCAGGCG	[1059]

	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	
001_aEF1a_Desulfurococcus_mobilis	ATG6TTATGT	GG---CATCC	AACAGCCATA	GCGGTT----	-----	-----	-----	G	GATACACGCC	GGTGATACAT	GTCCACACAG	CCAGCGTAGC [1070]
002_aEF1a_Sulfolobus_solfataricus	ATAGTAGTAT	GG---CATCC	TACGGCCTTG	GCAAAAT----	-----	-----	-----	G	GCTATACACC	AGTGATTACG	GTTCATACG	CGAGTTG6GC [1064]
003_aEF1a_Thermoplasma_acidophilum	aTcGtGcTca	ac---CaTcC	gAgCgTcAta	gCtCtC----	-----	-----	-----	G	gAtAcAaGcC	gGtAtTcCaC	gTgCaCaCaG	cCcAgGtGgC [1046]
004_aEF1a_Halobacterium_sp.	GTGCGTATGC	AG---CACCC	GTGCGTATC	ACCGCC----	-----	-----	-----	G	GGTACACGCC	GGTCTTCCAC	GGCCACACCG	CACAGGTCGC [1037]
005_aEF1a_Methanocaldococcus_jannaschii	GTTGTCTTAC	AG---CACCC	AACAGCAATT	ACAAGTT----	-----	-----	-----	G	GTACACACAC	AGTCTTCCAC	GCACACACAG	CACAGGTTGC [1052]
006_EF1a_Acanthamoeba_culbertsoni	ATCATCCTCA	AC---CACCC	TGGTCAGATC	CACGCT----	-----	-----	-----	G	GCTATGCCCC	CGTGCTCGAC	TGCCACACTG	CCCACATTGC [1037]
007_EF1a_Neoparamoeba_branchipila	ATCGTCCTTA	AC---CACCC	TGGAGAGATC	C6GTCC----	-----	-----	-----	G	GATACACCCC	CGTGTGGAT	TGCCACACCG	CTCACATTGC [1022]
008_EF1a_Planoprotostelium_aurantium	ATCATCCTCA	AC---CACCC	GGGAGAGATC	CACAAC----	-----	-----	-----	G	GTACG6CCCC	AGTCCTCGAC	TGCCACACCG	CCCACATTGC [1019]
009_EF1a_Physarum_polycephalum	ATCATCCTCA	AC---CACCC	CGGAACAATC	CACGCC----	-----	-----	-----	G	GATATGCCCC	CGTGCTCGAC	TGCCACACTG	CCCACATTGC [1019]
010_EF1a_Amoebidium_parasiticum												[867]
011_EF1a_Capsaspora_owczaraki	ATCGTCCTCA	AC---CACCC	CGGCCAGATC	TCCAAC----	-----	-----	-----	G	GCTACGCCCC	AGTCCTCGAC	TGCCACACTG	CCCACATTGC [1100]
012_EF1a_Monosiga_ovata	ATCATCCTGA	AC---CACCC	CGGCCAGATC	TCCAAC----	-----	-----	-----	G	GCTACGCGCC	CGTGCTCGAT	TGCCACACCG	CCCACATTGC [1079]
013_EF1a_Salpingoeca_amphoridium	ATCATCATGA	AC---CACCC	CGGTAGATC	TCCAAC----	-----	-----	-----	G	GCTATGCCCC	CGTGCTCGAC	TGCCACACTG	CCCACATTGC [1082]
014_EF1a_Ceanorhabditis_elegans	ATCATCATGA	AC---CATCC	AGGACAGATC	TCCAAC----	-----	-----	-----	G	GATACACTCC	AGTTCTCGAT	TGCCACACCG	CTCACATTGC [1106]
015_EF1a_Drosophila_melanogaster	ATCGTGCTGA	AC---CACCC	CGGTAGATTT	GCCAAC----	-----	-----	-----	G	GCTACACCCC	CGTGTGGAT	TGCCACACCG	CTCACATTGC [1106]
016_EF1a_Ornithorhynchus_anatinus	ATTATTCTAA	AT---CACCC	TGGCAGATC	AGCGCT----	-----	-----	-----	G	GGTACTCCCC	TGTCATCGAC	TGCCACACAG	CCCACATTGC [1106]
017_EF1a_Xenopus_laevis	ATTATACTCA	AT---CACCC	AGGACAGATT	AGTGCT----	-----	-----	-----	G	GCTATGCTCC	AGTCCTGGAC	TGTCATACAG	CTCATATTGC [1106]
018_EF1a_Saccharomyces_cerevisiae	ATTGTTTTGA	AC---CATCC	AGGTCAAATC	TCTGCT----	-----	-----	-----	G	GTTACTCTCC	AGTTTTGGAT	TGTCACACTG	CTCACATTGC [1100]
019_EF1a_Neurospora_crassa	ATCGTTCCTCA	AC---CACCC	CGGTAGATC	GGTGCC----	-----	-----	-----	G	GCTACGCCCC	CGTCTCGAC	TGCCACACTG	CCCACATTGC [1103]
020_EF1a_Cryptococcus_neoformans	ATCGTCCTTA	AC---CACCC	TGGTCAGATC	GGTGCC----	-----	-----	-----	G	GTACACCCCC	CGTTCTCGAC	TGTCACACTG	CCCACATTGC [1100]
021_EF1a_Schizophyllum_communis	ATCGTTCCTCA	AC---CACCC	CGGTAGATC	GGCGCT----	-----	-----	-----	G	GCTACGCAAC	CGTTCTCGAT	TGCCACACCG	CCCACATTGC [1100]
022_EF1a_Monoblepharis_insignis	ATTGTGATGA	AC---CACCC	TGGTCAGATC	AGCAAC----	-----	-----	-----	G	GTTACTCGCC	TGTGCTTGAC	TGCCACACGG	CTCACATTGC [986]
023_EF1a_Rhizophlyctis_harderi	ATTGTGTTGA	AC---CACCC	TGGTCAAATT	TCTGCT----	-----	-----	-----	G	GATACTCTCC	AGTCCTTGAT	TGTCACACTG	CTCACATTGC [986]
024_EF1a_Harpochytrium_sp.	ATCATCCTCA	AC---CACCC	CGGTAGATC	GGTGCC----	-----	-----	-----	G	GATACGCCCC	GGTTTTGGAC	TGCCACACCG	CCCACATTGC [938]
025_EF1a_Absidia_glauca	ATTGTCTCTTA	AC---CATCC	TGGTCAAATC	GGTGCT----	-----	-----	-----	G	GTTACGCTCC	TGTCTTGGAT	TGCCACACTG	CTCACATTGC [1100]
026_EF1a_Smittium_simulii	ATTGTTCTTA	AC---CACCC	TGGTCAAATC	TCTAAC----	-----	-----	-----	G	GTTATGCTCC	AGTTCTTGAT	TGCCACACTG	CTCACATTGC [1073]
027_EF1a_Chara_australis	ATCATCATGA	AT---CACCC	TGGACAGATT	GGTAAT----	-----	-----	-----	G	GCTATGCACC	GGTGCTGGAT	TGCCACACGT	GGCATATTGC [1019]
028_EF1a_Physcomitrella_patens	ATCATCATGA	AC---CACCC	GGGACAGATC	GGGAAC----	-----	-----	-----	G	GTACGCAACC	AGTGCTGGAT	TGCCACACGT	CGCACATTGC [1070]
029_EF1a_Acetabularia_acetabulum	ATCATCATGA	AC---CACCC	TGGTAAATC	GGCAAC----	-----	-----	-----	G	GGTACCAACC	GGTGCTGGAC	TGCCACACCG	CCCATATTGC [1019]
030_EF1a_Chaetomorpha_coliformis	ATCATTATGA	AC---CACCC	TGGCAGATTT	TCCAAC----	-----	-----	-----	G	GATACACCCC	CGTGCTGGAC	TGCCACACTG	CCCACATTGC [869]
031_EF1a_Arabidopsis_thaliana	ATCATCATGA	AC---CACCC	TGGTCAGATT	GGTAAC----	-----	-----	-----	G	GTTACGCCCC	AGTTCTCGAT	TGCCACACCT	CTCACATTGC [1070]
032_EF1a_Oryza_sativa	ATCATCATGA	AC---CACCC	TGGCAGATC	GGCAAC----	-----	-----	-----	G	GCTACGCCCC	AGTGCTGGAC	TGCCACACCT	CCCATATTGC [1070]
033_EF1a_Cyanidioschyzon_merolae	ATCATTCTGA	AT---CATCC	GGGCGAGATC	CGGGCC----	-----	-----	-----	G	GGTACGCGCC	CGTTGTTGAT	TGCCACTACTG	CCCATATCGC [1082]
034_EF1a_Porphyrha_jezoensis	ATCATCCTGA	AC---CACCC	TGGTCAGATC	CACGCT----	-----	-----	-----	G	GCTACGCGCC	GGTGCTGGAT	TGCCACACGG	CGCACATTGC [1082]
035_EF1a_Plasmodium_falciparum	ATTATATTGA	AC---CACCC	TGGTGAATTC	AAAAAT----	-----	-----	-----	G	GTATACACCC	AGTTTTGGAT	TGTCACACAT	CTCACATTTC [1064]
036_EF1a_Cryptosporidium_parvum	ATCGTCCTCA	AC---CACCC	AGGTGAATTC	AAGAAC----	-----	-----	-----	G	GTTACTCTCC	AGTGCTTGAC	TGTCACACCG	CTCACATTTC [1058]
037_EF1a_Tetrahymena_pyriformis	ATCATCATGA	AT---CACCC	CGGTGAATTC	CAAGCT----	-----	-----	-----	G	GTACACCCC	CGTTTTGGAT	TGCCACACTG	CTCACATTGC [1073]
038_EF1a_Stentor_coeruleus	ATTGTCTTGA	AT---CACCC	TGGCCAAATC	TCTGCT----	-----	-----	-----	G	GTTATACCCC	AGTTCTTGAT	TGCCACACTG	CTCACATTGC [1025]
039_EF1a_Phytophthora_parasitica	ATTGTGCTGA	AC---CACCC	CGGCCAGATC	GGCAAC----	-----	-----	-----	G	GCTACTCGCC	CGTGCTTGAC	TGCCACACGG	CGCAGTTGCG [1070]
040_EF1a_Mallomonas_rasilis	ATTATAATGT	CT---CACCC	TGGACAGATC	CAGAAC----	-----	-----	-----	G	GTTACTCTCC	TGTCTTGGAT	TGCCACACCT	CACACATTGC [1031]
041_EF1a_Phaeodactylum_tricornutum	ATTGTCTATGA	AC---CATCC	CGGTAGATC	TCCAAC----	-----	-----	-----	G	GATACTCTCC	CGTTCTTGAT	TGCCACACGG	CTCACATTGC [1070]
042_EF1a_Thalassiosira_pseudonana	ATCATCATGA	AT---CATCC	TGGGAAGATA	TTCTGT----	-----	-----	-----	G	GATACACTCC	AGTCCTTGAC	TGTCATACGG	CTCACATTGC [1070]
043_EF1a_Malawimonas_jakobiformis	ATCATCTTGA	AC---CACCC	CGGTAGATC	GGTGCC----	-----	-----	-----	G	GCTACACCCC	CGTGCTCGAT	TGCCACACCG	CCCACATTGC [1001]
044_EF1a_Malawimonas_californiana	ATTGTGCTGA	AC---CACCC	CGGCCAGATC	GGTGCC----	-----	-----	-----	G	GCTACACCCC	CGTGCTCGAC	TGCCACACCG	CGCACATTGC [1070]
045_EF1a_Trimastix_pyriformis	ATGATCATGA	AC---CACCC	TGGTCAGATC	CACTCC----	-----	-----	-----	A	ACTACACCCC	CGTTCTTGAC	TGCCACACCT	CCCACATTGC [1028]
046_EF1a_Dinenympha_exilis	ATTGTGATGG	CA---CACCC	TGGACAGATT	CAGAAT----	-----	-----	-----	G	GATACACACC	TGTGTTGGAT	TGCCACACGT	CACATATTGC [1010]
047_EF1a_Monocercomonoides_sp.	ATTGTCTATGT	CT---CACCC	AGGTCAAATC	TCAAAT----	-----	-----	-----	G	GATACACTCC	AGTTTTGGAT	TGCCACACAG	CACACATTGC [1070]
048_EF1a_Streblospio_trix	ATTGTTATGC	AA---CATCC	TGGACAAATC	CAAAAT----	-----	-----	-----	G	GATATACACC	AGTGCTTGAT	TGCCACACAT	CACACATTGC [1019]
049_EF1a_Histomonas_meleagridis	ATCATCTCCA	AC---CACCC	AGGTAAAGATC	CACGCT----	-----	-----	-----	G	GCTACCAACC	AGTTTTCGAC	TGCCACACAG	CTCACATTGC [1070]
050_EF1a_Trichomonas_vaginalis	ATCATCTCCA	AC---CACCC	AGGCAAGATC	CACGCC----	-----	-----	-----	G	GCTACCAACC	AGTTTTCGAC	TGCCACACAG	CCCACATTGC [1070]
051_EF1a_Hexamita_inflata	ATCATCATGA	AC---CACCC	CAAGAAGATC	AACCCA----	-----	-----	-----	G	GCTACACCCC	AGTCTCGAT	TGCCACACTT	CTCACATTGC [1022]
052_EF1a_Spiroplasma_vortens	ATCATCATGA	AC---CACCC	CAAGAAGATC	CAGCCG----	-----	-----	-----	G	GTTACACCCC	AGTGCTCGAC	TGCCACACCA	ACCACATTGC [1025]
053_EF1a_Andalucia_incarcerata	ATCATTTTGA	AC---CATCC	TGGTCAGATT	GGCGCT----	-----	-----	-----	G	GTTACTGCCC	CGTCTCGAT	TGCCACACTG	CTCACATTGC [1019]
054_EF1a_Seculamonas_ecuadoriensis	ATTGTGCTGA	AC---CATCC	GGGTAGATT	GGCCAG----	-----	-----	-----	G	GCTACTGCCC	GGTGCTCGAT	TGCCACACCT	CGCACATTGC [1070]
055_EF1a_Reclinomonas_americana	ATCATCCTGA	AC---CACCC	TGGCAGATC	GCTGCC----	-----	-----	-----	G	GCTACACCCC	GGTGCTGGAT	TGCCACACTG	CCCACATTGC [1070]
056_EF1a_Histiona_roides	ATCATCCTCA	AC---CACCC	TGGTCAAATC	GGCGCT----	-----	-----	-----	G	GTTACACCCC	AGTTGTTGGAT	TGCCACACTG	CTCACATTGC [1070]
057_EF1a_Jakoba_libera	ATCATCCTGA	AC---CACCC	TGGCAGATC	AACGCC----	-----	-----	-----	G	GATACGCCCC	CGTGTGGAT	TGCCACACCG	CCCACATTGC [1070]
058_EF1a_Stachyamoeba_lipophora	ATCATCATGA	AC---CACCC	TGGTGAATTC	CGTAAC----	-----	-----	-----	G	GATACAGCCC	AGTCATTGAT	TGCCACACTG	CTCACATTGC [1067]
059_EF1a_Naegleria_gruberi	ATTATCATGA	AC---CACCC	TGGTGAATTC	AGAAAT----	-----	-----	-----	G	GTTACTCTCC	AGTTTTGGAT	TGTCACACTT	CCCACATTGC [1022]
060_EF1a_Naegleria_andersoni	ATCATCATGA	AC---CACCC	TGGTGAATTC	AGAGCC----	-----	-----	-----	G	GTTACTCTCC	AGTTTTGGAT	TGTCACACCT	CCCACATTGC [1040]
061_EF1a_Acrasis_rosea	ATCATCTTGA	AC---CACCC	TGGTGAATTC	CAAAAC----	-----	-----	-----	G	GATACGCCCC	AGTTTTGGAT	TGCCACACTG	CTCACATTGC [1019]
062_EF1a_Sawyeria_marylandensis	ATTATTTTAA	AT---CACCC	AGGTAAATTC	TGCAAA----	-----	-----	-----	G	GTTACTCAAC	AGTTGTTGAT	TGTCATACAG	CATCATATTGC [1070]
063_EF1a_Bodo_saltans	ATCATCCTGA	AC---CACCC	CGGTAGATC	GGTAAC----	-----	-----	-----	G	GTTACTCTCC	CGTGCTCGAT	TGCCACACCA	GGCACATCGC [1010]
064_EF1a_Rhynchobodo_sp.	ATCGTCCTCA	AC---CACCC	CGGTAGATC	GGTAAC----	-----	-----	-----	G	GTTACTGCCCC	AGTGCTCGAC	TGCCACACCT	CCCACATTGC [1004]
065_EF1a_Rhynchomonas_nasuta	ATCGTGCTCA	AC---CACCC	GGGCCAGATC	TCCAAC----	-----	-----	-----	G	GCTACGCGCC	CGTGCTCGAC	TGCCACACGT	CGCACATTGC [1067]
066_EF1a_Dimastigella_trypaniformis	ATCGTCATGA	AC---CACCC	GGGCCAGATC	TGCAAC----	-----	-----	-----	G	GCTACGCGCC	CGTTCTGGAC	TGCCACACCG	CGCACATTGC [1025]

	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	
067_EF1a_Perkinsiella_amoebae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[915]
068_EF1a_Leishmania_major	ATCGTGCTGA	AC--CAACC	CGGCCAGATC	AGCAAC---	-----	-----	-G	GCTATGCGCC	GGTGTGGAC	TGCCACACGA	GCCACATTGC	[1070]
069_EF1a_Leishmania_infantum	ATCGTGCTGA	AC--CAACC	CGGCCAGATC	AGCAAC---	-----	-----	-G	GCTACGCGCC	GGTGTGGAC	TGCCACACGA	GCCACATTGC	[1070]
070_EF1a_Leishmania_braziliensis	ATCGTGCTGA	AC--CAACC	CGGCCAGATC	AGCAAC---	-----	-----	-G	GCTACGCGCC	GGTGTGGAC	TGCCACACGA	GCCACATTGC	[1070]
071_EF1a_Leishmania_tarentolae	ATCGTGCTGA	AC--CAACC	CGGCCAGATC	AGCAAC---	-----	-----	-G	GC	-----	-----	-----	[1032]
072_EF1a "L. tarentolae" FJ807257	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[915]
073_EF1a_Leptomonas_costaricensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[594]
074_EF1a_Leptomonas_bifurcata	ATCGTGCTGA	AC--CAACC	TGGCCAGATC	AGCAAC---	-----	-----	-G	GCTACGCGCC	NGTGTGGAC	TGCCACACCG	CCCACATCGC	[830]
075_EF1a_Leptomonas_podlipaevi	ATCGTGCTGA	AC--CAACC	CGGCCAGATC	AGCAAC---	-----	-----	-G	GCTACGCGCC	GGTGTGGAT	TGCCACACCG	CCCACA----	[814]
076_EF1a_Herpetomonas_muscarum	ATCATCTGA	AC--CAACC	CGGCCAGATC	AGCAAC---	-----	-----	-G	GCTACGCCCC	CGTCTGGAT	TGCCACACCA	GCCACATCGC	[983]
077_EF1a_Herpetomonas_pessoai	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[771]
078_EF1a_Blastocirithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[766]
079_EF1a_Trypanosoma_brucei	ATCATCTGA	AC--CAACC	CGGACAGATT	GGAAAC---	-----	-----	-G	GTTATGCGCC	CGTGTGGAC	TGCCACACAT	GCCACATTGC	[1070]
080_EF1a_T. brucei equiperdum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[774]
081_EF1a_Trypanosoma_aviium	ATCATCTGA	AC--CAACC	TGGCCAGATC	GGCAAC---	-----	-----	-G	GCTATGCCCC	CGTGTGGAC	TGCCACACCT	GCCACATTGC	[1016]
082_EF1a_Trypanosoma_cruzi	ATCATCTGA	AC--CAACC	CGGCCAGATC	GGCAAC---	-----	-----	-G	GCTATGCGCC	GGTGTGGAT	TGCCACACCT	GCCACATCGC	[1070]
083_EF1a_Trypanosoma_rangeli	ATTATCTGA	AC--CAACC	CGGCCAGATC	GGCAAC---	-----	-----	-G	GCTACGCGCC	GGTGTGGAT	TGCCACACCT	GCCACATCGC	[1070]
084_EF1a_Ploetia_costatum	ATTATCTGA	AT--CAACC	TGGTCAGATC	GGAAAC---	-----	-----	-G	GTTACGCAAC	AGTGTGGAC	TGCCACACTT	GCCACATTGC	[1067]
085_EF1a_Ploetia_oblonga_cDNA	ATTATCTGA	AT--CAACC	TGGTCAGATC	GGAAAC---	-----	-----	-G	GTTACGCAAC	AGTGTGGAC	TGCCACACTT	GCCACATTGC	[1067]
086_EF1a_Ploetia_oblonga_gDNA	ATCATCTGA	AC--CAACC	CGGTAGATC	TCCAAC---	-----	-----	-G	GGTACGCCCC	CGTCTGGAT	TGCCACACGT	GCCATATTGC	[1067]
087_EF1a_Entosiphon_sp.	ATTGTCTGA	AC--CAACC	CGGTAGATC	TCCAAC---	-----	-----	-G	GCTACGCCCC	GGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
088_EF1a_Entosiphon_sulcatum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[915]
089_EF1a_Peranema_tricophorum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[938]
090_EF1a_P. tricophorum FJ807243	ATTGTCTGA	AC--CAACC	TGGTCAAAAT	CAAAAC---	-----	-----	-G	GATACGCCCC	AGTGTGGAT	TGCCACACCT	CACACATTGC	[1061]
091_EF1a_Colacium_elongatum	ATTATCTGA	AC--CAACC	TGGTCAGATT	GGCAAC---	-----	-----	-G	GTTATGCCCC	CGTCTGGAT	TGCCACACGT	GCCACATTGC	[1067]
092_EF1a_Colacium_mucronatum	ATTATCTGA	AC--CAACC	TGGTCAGATT	GGCAAT---	-----	-----	-G	GCTATGCCCC	TGTTTGGAT	TGCCACACTT	GCCACATCGC	[1013]
093_EF1a_Colacium_sp.	ATCATCTGA	AC--CAACC	CGGCCAGATC	GGCAAC---	-----	-----	-G	GTTATGCCCC	CGTCTGGAT	TGCCACACTT	GCCACATTGC	[1067]
094_EF1a_Colacium_vesiculosum	ATCATCTGA	AC--CAACC	TGGTCAGATT	GGCAAT---	-----	-----	-G	GCTATGCCCC	TGTTTGGAT	TGCCACACTT	GCCACATCGC	[1025]
095_EF1a_Cryptoglena_pigra	ATTATCTGA	AC--CAACC	TGGCCAGATT	GGCAAC---	-----	-----	-G	GCTACGCAAC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
096_EF1a_Cryptoglena_skujai	ATCATCTGA	AC--CAACC	TGGCCAGATT	GGTAAT---	-----	-----	-G	GCTACGCAAC	CGTGTGGAT	TGCCACACTT	GTCACTATGC	[1067]
097_EF1a_Discoplistis_adunca	ATCATCTGA	AC--CAACC	TGGCCAAAT	GGGAAT---	-----	-----	-G	GCTACGCGCC	AGTGTGGAT	TGCCACACGT	GCCACATTGC	[1067]
098_EF1a_Euglena_clara	ATTATCTGA	AC--CAACC	TGGCCAGATT	GGCAAT---	-----	-----	-G	GCTACGCCCC	TGTGTGGAT	TGCCACACCT	GCCACATCGC	[947]
099_EF1a_Euglena_deses	ATCATCTGA	AT--CATCC	AGGCCAGATT	GGAAAT---	-----	-----	-G	GATATGCTCC	AGTATTGGAT	TGCCACACTT	GCCACATCGC	[1067]
100_EF1a_Euglena_exilis	ATCATCTGA	AC--CAACC	AGGTAGATT	GGCAAC---	-----	-----	-G	GCTATGCTCC	AGTGTGGAC	TGCCACACCT	GCCACATTGC	[947]
101_EF1a_Euglena_geniculata	ATTATCTGA	AT--CAACC	TGGTCAGATC	GGAAAC---	-----	-----	-G	GTTACGCAAC	AGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
102_EF1a_Euglena_gracilis	ATCATCTGA	AC--CAACC	TGGTCAGATT	GGCAAC---	-----	-----	-G	GTTATGCGCC	CGTGTGGAC	TGCCACACTT	GCCACATTGC	[1067]
103_EF1a E. gracilis_X16890	ATCATCTGA	AC--CATCC	TGGCCAGATT	GGGAAC---	-----	-----	-G	GGTACGCAAC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1070]
104_EF1a E. gracilis_FJ807238	ATCATCTGA	AC--CAACC	TGGTCAGATT	GGCAAC---	-----	-----	-G	GTTATGCGCC	CGTGTGGAC	TGCCACACTT	GCCACATTGC	[1070]
105_EF1a_Euglena_granulata	ATCATCTGA	AT--CAACC	TGGTCAGATT	GGCAAC---	-----	-----	-G	GCTATGCAAC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
106_EF1a_Euglena_hiemalis	ATCGTCTGA	AC--CAACC	TGGCCAGATC	GGAAAC---	-----	-----	-G	GCTATGCAAC	AGTCTGGAT	TGCCACACTT	GCCACATTGC	[1067]
107_EF1a_Euglena_longa	ATTATCTGA	AC--CAACC	TGGCCAGATT	GGAAAT---	-----	-----	-G	GTTATGCAAC	AGTCTGGAT	TGCCACACTT	GCCATATTGC	[584]
108_EF1a_Euglena_stellata	ATCATTTGA	AC--CATCC	TGGGCAAAAT	GGCAAC---	-----	-----	-G	GTTATGCAAC	TGTGTGGAT	TGTACACAT	GCCACATTGC	[947]
109_EF1a_Euglena_viridis	ATCATCTGA	AC--CATCC	TGGCCAGATT	GGCAAT---	-----	-----	-G	GTTATGCCCC	CGTCTGGAT	TGCCACACTT	GCCACATTGC	[1025]
110_EF1a_Eutreptia_pertyi	ATCATCTGA	AT--CAACC	TGGTCAGATT	GGCAAC---	-----	-----	-G	GCTATGCAAC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
111_EF1a_Eutreptiella_pomquetensis	ATTATCTGA	AT--CAACC	TGGTCAGATC	GGAAAC---	-----	-----	-A	GTTACGCAAC	AGTGTGGAC	TGCCGCACTT	GCCACATTGC	[677]
112_EF1a_Lepocinclis_acus_var._major	ATCATCTGA	AC--CAACC	TGGCCAGATT	GGAAAT---	-----	-----	-G	GTTATGCCCC	AGTCTGGAT	TGCCACACTT	GCCACATTGC	[1067]
113_EF1a_Lepocinclis_ovum	ATCATCTGA	AC--CAACC	TGGCCAGATT	GGGAAT---	-----	-----	-G	GTTATGCCCC	AGTCTGGAT	TGCCACACTT	GCCACATTGC	[947]
114_EF1a_Lepocinclis_spyrogyroides_clone_38	ATCATCTGA	AC--CAACC	TGGCCAGATT	GGGAAT---	-----	-----	-G	GTTATGCTCC	CGTGTGGAT	TGCCACACTT	GCCACATTGC	[380]
115_EF1a L. spyrogyroides_clones_1_&_4	ATTATCTGA	AC--CAACC	TGGCCAGATT	GGCAAC---	-----	-----	-G	GCTACGCCCC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
116_EF1a_Lepocinclis_steinii	ATCATCTGA	AC--CAACC	TGGTCAGATT	GGCAAC---	-----	-----	-G	GTTATGCCCC	TGTGTGGAT	TGTACACAT	GCCACATTGC	[947]
117_EF1a_Lepocinclis_tripteris	ATTATCTGA	AC--CAACC	TGGGCAAAAT	GGAAAT---	-----	-----	-G	GTTATGCCCC	TGTGTGGAT	TGCCACTT	GCCATATCGC	[1067]
118_EF1a_Monomorphina_aenigmatica	ATCATCTGA	AC--CAACC	TGGCCAGATT	GGAAAC---	-----	-----	-G	GCTATGCCCC	CGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
119_EF1a_Monomorphina_pyrum	ATTATCTGA	AC--CAACC	TGGCCAGATT	GGCAAC---	-----	-----	-G	GCTACGCCCC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
120_EF1a_Monomorphina_sp.	ATCATCTGA	AC--CATCC	TGGCCAGATT	GGCAAT---	-----	-----	-G	GCTATGCCCC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
121_EF1a_Phacus_orbicularis	ATCGTCTGA	AT--CATCC	TGGCCAGATT	GGGAAC---	-----	-----	-G	GTTACGCAAC	AGTGTGGAT	TGCCACACTT	GCCACATTGC	[1025]
122_EF1a_Phacus_similis	ATCATCTGA	AC--CATCC	TGGCCAAATC	TCCAAT---	-----	-----	-G	GCTATGCCCC	CGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
123_EF1a_Strombomonas_acuminata	ATCATCTGA	AC--CAACC	TGGCCAGATT	GGCAAC---	-----	-----	-G	GCTATGCCCC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1070]
124_EF1a_Strombomonas_borsteniensis	ATCATCTGA	AC--CATCC	TGGTCAGATT	GGCAAC---	-----	-----	-G	GCTATGCCCC	AGTGTGGAT	TGCCACACTT	GCCACATTGC	[947]
125_EF1a_Strombomonas_napiformis	ATCATCTGA	AC--CAACC	TGGCCAGATT	GGAAAT---	-----	-----	-G	GCTATGCAAC	AGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
126_EF1a_Strombomonas_ovalis	ATCATCTGA	AC--CAACC	TGGTCAGATT	GGCAAC---	-----	-----	-G	GCTATGCCCC	AGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
127_EF1a_Strombomonas_triquetra	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[891]
128_EF1a_Trachelomonas_conradi	ATCATCTGA	AC--CATCC	TGGCCAGATT	GGAAAT---	-----	-----	-G	GCTATGCCCC	CGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
129_EF1a_Trachelomonas_echinata	ATCATCTGA	AC--CAACC	TGGCCAGATT	GGAAAT---	-----	-----	-G	GGTACGCAAC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
130_EF1a_Trachelomonas_lefevrei	ATCATCTGA	AC--CAACC	AGGCCAAAT	GGGAAC---	-----	-----	-G	GCTATGCCCC	TGTGTGGAT	TGCCACACTT	GCCACATTGC	[1067]
131_EF1a_Trachelomonas_plantonica	ATCATCTGA	AC--CATCC	AGGCCAGATT	GGGAAT---	-----	-----	-G	GCTATGCCCC	CGTCTGGAT	TGCCACACTT	GCCACATTGC	[947]
132_EF1a_Trachelomonas_similis	ATCATCTGA	AC--CAACC	TGGCCAAAT	GGCAAT---	-----	-----	-G	GATATGCAAC	AGTATTGGAT	TGCCACACGT	GCCACATTGC	[947]

	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	
133_EFL1a_Trachelomonas_volvocina	ATCATTCTGA	AT---CACCC	TGGTCAGATT	GGCAAT----	-----	-----	-----G	GATATGCACC	AGTTCTGGAT	TGCCACACCT	GCCACATTGC	[1067]
134_EFLenvironmental_sample_EU810232	AAGGTCCTCA	CC---CACCC	CAACGAGCTC	AAGTGC----	-----	-----	-----G	GATACTGCCC	CATCGCCTTC	ATCAGGACCG	CCAGGCTGTC	[1160]
135_EFLCreolimax_fragrantissima												[994]
136_EFLSphaeroforma_arctica	CAGGTTATGA	AC---CACCC	TGGTGTACTA	AAGGCA----	-----	-----	-----G	GATACTGTCC	CATTGCCTTT	GTCCGTAACC	GACGTTCCGC	[1160]
137_EFLMonosiga_brevicollis	CAGATCATGA	AC---CACCC	CGGAGAGCTC	AAGGTT----	-----	-----	-----G	GCTACTGCCC	CATTGCCTTC	GTCCGCACCT	CCCGCTCGGC	[1160]
138_EFLXenopus_laevis	CAGATCATGA	AC---CACCC	CGGTGAGCTC	AAGGTC----	-----	-----	-----G	GCTACTGCCC	GGTCGCCTTC	GTCCGCACGG	GCCGCTCCGC	[1160]
139_EFLSpizellomyces_punctatus	CAGACTCTCG	AC---ATCCC	CGGTGAGGTC	AAGCCC----	-----	-----	-----G	GATACTCTCC	CATTGGTTTC	GTCCGATGCG	GTGCTCTGTC	[1166]
140_EFLGaertneriomyces_semiglobiferous												[506]
141_EFLAllomyces_macrogynus	CAGACGCTCG	AT---ATCCC	CTCGGAGGTC	AAGCCC----	-----	-----	-----G	GCTACTCGCC	CATCGGCTTC	GTCCGCTGCG	GCCGCTCGGC	[1166]
142_EFLBlastocladiella_emersonii	CAGACGCTCG	AC---ATCCC	CTCGGAGGTC	AAGCCC----	-----	-----	-----G	GCTACTCGCC	CATCGGCTTC	GTCCGCTGCG	GTGCTCTGCG	[1166]
143_EFLCatenophylctis_sp.												[506]
144_EFLCoelomomyces_stegomyiae												[506]
145_EFLEntomophthora_muscae												[503]
146_EFLBasidiobolus_ranarum												[503]
147_EFLConidiobolus_coronatus												[503]
148_EFLTetraselmis_tetrathele	CAGACCTTGG	AC---ATCCC	CGGTGAGATC	AAGGTC----	-----	-----	-----G	GCTACTCCCC	CATCGGCTTC	GTCCGCTGCG	GCCGCGCTGC	[1109]
149_EFLMesostigma_viride												[879]
150_EFLOstreococcus_tauri	CAAACTTTGG	AC---ATCCC	GGGTGAACCT	AAGGTT----	-----	-----	-----G	GCTACTCCCC	GATCGGTTTC	GTCCGTTGCG	GTGCTCCGCG	[1169]
151_EFLNephroselmis_olivacea												[957]
152_EFLMicromonas_pusilla	CAGGTTCTCG	AC---GTCCC	CGGCCAGCTC	AAGTGC----	-----	-----	-----G	GCTACTCCCC	CGTCGCCTTC	GTCCGCTGCG	CCCGCTCCGC	[1118]
153_EFLParachlorella_kessleri												[942]
154_EFLHelicosporidium_sp.	CAGACCTTGG	AC---ATCCC	CGGCCAGGTC	AAGGCC----	-----	-----	-----G	GCTACAGCCC	CATCGGCTTC	GTCCGCTGCG	GCCGCGCCGC	[1181]
155_EFLGonium_pectorale	CAGACCTTGG	AT---ATCCC	CGGCCAGGTC	AAGAAG----	-----	-----	-----G	GCTACTCGCC	CATCGGCTTC	GTCCGCTGCG	GCCGCTCCGC	[1163]
156_EFLChlamydomonas_reinhardtii	CAGACCTTGG	AC---ATCCC	CGGTGAGGTC	AAGAAG----	-----	-----	-----G	GCTACTCGCC	CATCGGCTTC	GTCCGCTGCG	GCCGCTCCGC	[1163]
157_EFLScenedesmus_obliquus	CAGACCTTGG	AC---ATCCC	AGGCGAGGTC	AAGGCC----	-----	-----	-----G	GCTACTCCCC	CATCGGCTTC	GTCCGCTGCG	CCCGCTCCGC	[1163]
158_EFL"Oryza_sativa"	CAGACCTTGG	AC---ATCCC	AGGTGAGGTC	AAGAAG----	-----	-----	-----G	GCTACAGCCC	CATCGGCTTC	GTCCGCTGCG	GTCCGCTGCG	[1163]
159_EFLChlorococcum_sp.	CAGACTCTGG	AC---ATCCC	TGGTGAGCTT	AAGGTG----	-----	-----	-----G	GTACTCTCCC	GATTGGCTTC	GTCCGCTGCG	CCCGCTGCTG	[1112]
160_EFLOchlochaete_hystrix	CAGACCTTGG	AC---ATCCC	GGGTGAGGTC	AAGGCA----	-----	-----	-----G	GTACTCTCCC	CATTGGCTTC	GTCCGCTGCG	GCCGCTCCGC	[1121]
161_EFLUlva_intestinalis	CAGACCTTGG	AC---ATCCC	CGGTGAGGTC	AAGGCC----	-----	-----	-----G	GCTACTCGCC	GATTGGCTTC	GTCCGCTGCG	CCCGCTGCTG	[1112]
162_EFLUrospora_sp.	CAGACGCTTG	AC---ATCCC	CGGTGAGCTC	AAGGTT----	-----	-----	-----G	GGTACTCGCC	CATTGGCTTC	GTCCGCTGCG	GCCGCTCCGC	[965]
163_EFLGymnochlorella_stellata_ppc	CAAGTATTGG	AC---GTGCC	TCATGAAGTC	AAGGCT----	-----	-----	-----G	GTATTCTACC	CATCGGATTC	GTCCGATGCG	GTCCGCTGCG	[1199]
164_EFLLotharella_amoebiformis_ppc	CAGGTGCTGG	AC---GTCCC	CCATGAGCTC	AAGCCT----	-----	-----	-----G	GCTACTCGCC	GATTGGCTTC	GTCCGCTGCG	GACGCTGCGC	[944]
165_EFLChlorarachnion_reptans_ppc	CAAGTCTTGG	AT---GTTCC	ACACGAGATC	AAGCCG----	-----	-----	-----G	GATACTCCCC	AATCGGTTTC	GTAAAGTGCG	GACGGTCAGC	[971]
166_EFLChondrus_crispus	CAGATCTCTG	AC---CATCC	TTTGTAATTC	AAGATT----	-----	-----	-----G	GCTACGCCCC	AGTGGCATTC	GTCCGCACCT	CCCGATCTGC	[1160]
167_EFLGracilaria_changii	CAGATTCTGG	AT---CATCC	TGGTCAGCTC	AAGGTT----	-----	-----	-----G	GATACGCCCC	AGTTGGCTTC	GTCCGCACCT	GATGGTCTGC	[1160]
168_EFLCryptomonas_ovata	CAGACTCTGG	ACAACATCCC	CGGAGAGCTC	AAGGTT----	-----	-----	-----G	GGTATTCTCC	GATCGGTTTC	GTCCGATGCG	GCCGCGGACG	[1121]
169_EFLGuillardia_theta	CAGACCTTGG	ACAACATCCC	CGGAGAGCTT	AAGACA----	-----	-----	-----G	GATACTCACC	AATTGGCTTC	GTCCGATGCG	GACGTGCTGC	[1166]
170_EFLRhodomonas_salina	CAGACTCTCG	ACAACATCCC	GGGAGAGCTC	AAGGTC----	-----	-----	-----G	GATACTCCCC	CATTGGCTTC	GTCCGCTGCG	GCCGCGGACG	[1163]
171_EFLPavlova_lutheri	CAGACGCTCG	AC---ATCCC	GGGCGAGCTC	AAGGTG----	-----	-----	-----G	GCTACTCGCC	CATCGGCTTC	GTCCGCTGCG	GCCGCGCCGC	[1160]
172_EFLChrysochromulina_sp.	CAGACCTTGG	AC---ATCCC	CGGCGAGATC	AAGCTC----	-----	-----	-----G	GCTACTCGCC	GATCGGCTTC	GTCCGCTGCG	GCCGCTCCGC	[980]
173_EFLIsochrysis_galbana	CAGACCTTGG	AC---ATCCC	CGGCGAGATC	AAGATC----	-----	-----	-----G	GCTACTCTCC	GATTGGCTTC	GTCCGCTGCG	GCCGCGCCGC	[1163]
174_EFLEmiliana_huxleyi	CAGACCTTGG	AC---ATCCC	CGGCGAGATC	AAGCTG----	-----	-----	-----G	GCTACTCGCC	GATCGGCTTC	GTCCGCTGCG	GCCGCTCCGC	[1169]
175_EFLAchnanthes_kuwaitensis	GCCGTCACAG	AG---CACCC	TGGTGTGCTT	AAGCCC----	-----	-----	-----G	GATACTGCCC	GTAGGTTTTC	GTCCGTAACC	CCAAGGTCGC	[1070]
176_EFLAsterionella_glacialis	GCCGTTCAAG	AG---CACCC	TGGTGTGCTC	AAGCCT----	-----	-----	-----G	GATACTGCCC	ACTTATCTTC	TCCCGTACCG	CCAAGGTTGC	[1070]
177_EFLDetonula_confervacea	GCCGTGCAGG	AG---CACCC	TGGTGTGCTC	AAGGCT----	-----	-----	-----G	GATACGCGCC	GTGTTGCTTC	GTCCGTAACC	CTAAGGTTGC	[1070]
178_EFLThalassiosira_pseudonana	GCCGTGCAGG	AG---CACCC	CGGTGTGCTC	AAGCCC----	-----	-----	-----G	GATATGACCC	CGTTGTTTTC	GTCCGTAACC	CCAAGGTTGC	[1133]
179_EFLPerkinsus_marinus	CAAGTTTTGG	AC---GTCCC	AGGCGAGATT	AAGGTT----	-----	-----	-----G	GTATTCTTCC	TATCGGTTTC	GTCCGTAACC	GCCGCTCTGC	[1178]
180_EFLOxyrrhis_marina	CAGGTGTTGG	AC---ATCCC	CGGTGAGATC	AAGATG----	-----	-----	-----G	GGTATTCCCC	GATTGGTTTC	GTCCGCTGCG	GCCGCGCCGC	[1172]
181_EFLKarlodinium_micrum	CAAGTTTTGG	AC---ATTCC	TAATGAATTC	AAGGTC----	-----	-----	-----G	GGTACTCTCC	GATCGGATTC	GTCCGTTGTC	GCCGCTCAGC	[1166]
182_EFLAmphidinium_carterae	CAGGTCCTGG	AC---ATCCC	CAATGAGATC	AAGGTC----	-----	-----	-----G	GCTACTCTCC	CATCGGCTTC	GTCCGCTGTC	GCCGCTGCTG	[1166]
183_EFLHeterocapsa_triquetra	CAGGTCCTGG	AC---ATCCC	GAACGAGATC	AAGACG----	-----	-----	-----G	GCTACTCGCC	CATTGGCTTC	GTCCGCTGCG	GCCGCGCTGC	[1166]
184_EFLPlanoglabratella_operculis	TTTGTTCAGG	AT---CATCC	AGGCAAAATTA	GCCCATGCTA								[961]
185_EFLReticulomyxa_filosa	TTTGTTCAGG	AT---CATCC	AGGCAAAATTA	AAAGCCGCAA	CTAATGACGG	AAAAGGCAAC	TATAAAGGAG	GATTTACCCC	ATCGATTAC	ATTAGAAGTC	CTAAGGCTCC	[1106]
186_EFLLotharella_amoebiformis	AAGGTCACAG	AG---CACCC	CGGTGAGGTC	AAGGTC----	-----	-----	-----G	GATACTGTCC	TCATGTTTTC	GTCCGTAACC	CTAAGTCTGC	[1241]
187_EFLGymnochlorella_stellata	AAGATTCAAG	AG---CACCC	AGGAAAGCTC	AAAGTT----	-----	-----	-----G	GATACACTCC	TCTCGTCTTC	ATCCGTAACC	CTAAGGCTCC	[1214]
188_EFLChlorarachnion_reptans	AAGATTCAAG	AG---CACCC	TGGTAAAGTC	AAGGTT----	-----	-----	-----G	GGTACACCCC	GCTTGTGCTT	GTCCGTAACC	CTAAGGCTCC	[1145]
189_EFLDiplonema_ambulator	CAGGTGCTGG	AC---CACCC	TGGTGAAGTC	AAGGTC----	-----	-----	-----G	GGTACACCCC	CGTGTGCTTC	GTCCGTAACC	GTAGGTCCTG	[764]
190_EFLDiplonema_papillatum	CAGGTGCTGG	AC---CACCC	GGGTGAAGTC	AAGGTT----	-----	-----	-----G	GGTACACCCC	GGTGTGCTTC	GTCCGTAACC	GTAGATCGGC	[1136]
191_EFLRhynchopus_euleeides	CAGGTGCTGG	AC---CACCC	CGGTGAGGTC	AAGGTC----	-----	-----	-----G	GCTACACCCC	CGTGTGCTTC	GTCCGTAACC	GCCGCTCCGC	[779]
192_EFLPetalomonas_cantuscygni												[924]
193_EFLTrypanoplasma_borrel												[1017]
194_EFLNeobodo_salien	ATGGTCTCTG	AC---CACCC	GGGCCAGCTC	CGCGTT----	-----	-----	-----G	GCTACACCCC	GACGTGCTAC	GTCCGACCT	GACAGTGGC	[1133]

	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	
001_aEF1a_Desulfurococcus_mobilis	CTGCAGGATA	ACCGAGATCA	TA-----	-GCGAAGATT	GACCCGAGG-	-----AC	CGGTAAAGGAA	ATC-----	-----GAGAA	GAACCCGCAC	TTCTGAAGC	[1150]
002_aEF1a_Sulfolobus_solfataricus	ATGCAGAGTT	TCAGAATTAG	TA-----	-TCAAAACTA	GATCCAAGA-	-----AC	TGGTCAAGAA	GCA-----	-----GAAAA	GAATCCACAG	TTCTTGAAC	[1144]
003_aEF1a_Thermoplasma_acidophilum	aTgCaGgAta	gATgAGaTaG	tC-----	-AaGaCcCtT	aAcCcGaAg-	-----ga	tGgAaCaAcT	cTg-----	-----AaGgA	gAaAcCaGaC	tTcAtAaAgA	[1126]
004_aEF1a_Halobacterium_sp.	GTGTACCATC	AATTCCATCG	AG-----	-AAGAAGATG	GACCCGGCG-	-----TC	CGGGAGAGACC	CAG-----	-----GAGGA	GAACCCGGAG	TTATCCCAAT	[1117]
005_aEF1a_Methanocaldococcus_jannaschii	ATGTACATTTC	ATTGAGTTGT	TG-----	-AAGAAATTA	GACCCAAGA-	-----AC	AGGGCAAGATC	ATT-----	-----GAAGA	GAACCCACAG	TTCTTGAAG	[1132]
006_EF1a_Acanthamoeba_culbertsoni	CTGCAAGTTC	AAGGAGCTCC	TT-----	-GAGAAGGTC	GACCGTCGT-	-----TC	TGGTAAGGAAG	CTC-----	-----GAGGA	CAGCCCCAAG	GCCGTCAGT	[1117]
007_EF1a_Neoparamoeba_branchipila	CTGCAAGTTC	GCTGAGATCA	CC-----	-GAGAAGATC	GACCGTCGT-	-----TC	CGGTAAAGACC	GTC-----	-----GAGAA	GGAGCCCCAG	GCTGTCAAGT	[1102]
008_EF1a_Planoprotostelium_aurantium	CTGCAAGTTC	GCTGAGATCA	CC-----	-TCCAAGATC	GACAGACGT-	-----TC	CGGAAAGGAGC	GTC-----	-----GAGGC	TCTCCCCAAG	TCCGTCAGT	[1099]
009_EF1a_Physarum_polycephalum	CTGCAAGTTC	ACTGAGATCT	TG-----	-TCCAAGGTC	GACCGTCGT-	-----AC	TGGCAAGGCC	CTC-----	-----GAGGA	GGTCCCCAAG	AACATCAAGA	[1099]
010_EF1a_Amoebidium_parasiticum												[867]
011_EF1a_Capsaspora_owczaraki	CTGCAAGTTC	CAGGACATCA	AG-----	-GAGAAGTGC	GACCGTCGC-	-----TC	TGGCAAGAAG	CTC-----	-----GAAGA	CGCCCCCAAG	TTGTCAGAT	[1180]
012_EF1a_Monosiga_ovata	TTGCAAGTTC	AACGAGATCA	AG-----	-GAGAAGATT	GACCGTCGT-	-----TC	GGGCAAGAAG	ATC-----	-----GAGGA	CTTCCCCAAG	GCCGTCAGT	[1159]
013_EF1a_Salpingoeca_amphoridium	CTGCAAGTTC	CACGAGATCA	CC-----	-GCCAAGCTC	GACCGCCGC-	-----AC	GGGCAAGGCT	ACT-----	-----GAGGA	GAACCCCAAG	TTGTCAGAA	[1162]
014_EF1a_Ceanorhabditis_elegans	CTGCAAGTTC	AACGAGCTTA	AG-----	-GAGAAGGTT	GACCGTCGT-	-----AC	CGGTAAAGAAG	GTT-----	-----GAGGA	CTTCCCCAAG	TTCTTCAAGT	[1186]
015_EF1a_Drosophila_melanogaster	TTGCAAGTTC	GCTGAGATCA	AG-----	-GAGAAGGTC	GACCGTCGT-	-----TC	CGGCAAGACC	ACT-----	-----GAGGA	GAACCCCAAG	TTCATCAAGT	[1186]
016_EF1a_Ornithorhynchus_anatinus	CTGCAAGTTC	GCCGAGCTGA	AG-----	-GAGAAGATT	GACCGACGT-	-----TC	TGGCAAGAAG	CTG-----	-----GAGGA	GAGCCCAAG	TCCTTGAAT	[1186]
017_EF1a_Xenopus_laëvis	TTGTAAGTTT	GCTGAGCTGA	AG-----	-CAAAAGATT	GACAGAAGA-	-----AG	TGGCAAGAAG	CTG-----	-----GAAGA	CGATCCAAA	TTCTTGAAGT	[1186]
018_EF1a_Saccharomyces_cerevisiae	TTGTAGATT	GACGAATTGT	TG-----	-GAAAAGAAC	GACAGAAGA-	-----TC	TGGTAAGAAG	TTG-----	-----GAAGA	CCATCCAAAG	TTCTTGAAGT	[1180]
019_EF1a_Neurospora_crassa	CTGCAAGTTC	GCCGAGCTCC	TC-----	-GAGAAGATC	GACCGCCGT-	-----AC	TGGTAAGGCT	GTT-----	-----GAGGC	CTTCCCCAAG	TTCTTCAAGT	[1183]
020_EF1a_Cryptococcus_neoformans	CTGCAAGTTT	GCTGAGTTGA	TC-----	-GAGAAGATT	GACCGACGA-	-----AC	CGGTAAAGTTC	ATG-----	-----GAGGC	CGCCCCCAAG	TTGTCAGAT	[1180]
021_EF1a_Schizosaccharomyces_m commune	CTGCAAGTTC	GCCGAGCTCG	TC-----	-GAGAAGATC	GATCGCCGG-	-----AC	GGGTAAGTCTG	CTC-----	-----GAGGC	GAGCCCAAG	TTGTCAGAT	[1180]
022_EF1a_Monoblepharis_insignis	CTGCCGTTTC	AACAATATCC	TG-----	-CAGAAGATC	GACCGTCGT-	-----TC	GGGTAAAGTG	CTC-----	-----GAAGA	GAACCCTAAG	TTATCAAGT	[1066]
023_EF1a_Rhizophlyctis_harderi	TTGCAAAATC	GCTGAACCTG	TG-----	-GAAAAGTTG	GATCGTCGT-	-----AC	TTCAAAGAAA	CTT-----	-----GAAGA	CAATCCAAA	TTTATCAAGT	[1066]
024_EF1a_Harpochytrium_sp.	CTGCAAGTTC	GCCGAGCTTG	TT-----	-GAGAAGATT	GACCGTCGT-	-----TC	CGGAAAGAAG	ATG-----	-----GAGGA	CGGCCCCAAG	TTGTCAGAT	[1018]
025_EF1a_Absidia_glauca	CTGTAAGTTC	GCCGAACCTC	TG-----	-GAAAAGATC	GATCGTCGT-	-----TC	CGGTAAAGAG	CTC-----	-----GAAGA	TGCTCCCAAG	TTGTCAGAT	[1180]
026_EF1a_Smittium_simulii	ATGCAAAATC	ACTGAAATTA	TC-----	-GAAAAGATG	GATCGTCGT-	-----TC	CGGTAAAACT	CTT-----	-----GAAGC	TGCTCTCTAA	TTTATCAAGT	[1153]
027_EF1a_Chara_australis	GGTGAATTTT	GCTGAACCTG	TG-----	-GTGAAGATC	GACCGCCGA-	-----AC	GGGCAAGGAA	ACT-----	-----GAGAA	GGAGCCTAAG	TTTTTGAAGA	[1099]
028_EF1a_Physcomitrella_patens	GGTGAAGTTT	GCGGAGATCT	TG-----	-ACGAAGGTG	GACAGGCCGA-	-----TC	GGGTAAAGGAG	CTG-----	-----GAGAA	GGAGCCGAAG	TTTTTGAAGA	[1150]
029_EF1a_Acetabularia_acetabulum	TGTCAAGTTC	AAGGATATCC	AG-----	-AAGAAGGTG	GACAGAAGG-	-----TC	CGGAAAGGTTG	GTG-----	-----GAGGA	GTCCCCCAAG	TTATCAAAA	[1099]
030_EF1a_Chaetomorpha_coliformis	CGTCAAGTTC	GAACACTATCC	AG-----	-CAAAAGGTT	GACAGGCGT-	-----AC	CGGTAAAGTTC	ACT-----	-----GAGGA	GAATCCTAAG	TTATCAAAA	[949]
031_EF1a_Arabidopsis_thaliana	AGTCAAGTTC	TCTGAGATCT	TG-----	-ACCAAGATT	GACAGGCGT-	-----TC	TGGTAAGGAG	CTG-----	-----GAGAA	GGAGCCCAAG	TTTTTGAAGA	[1150]
032_EF1a_Oryza_sativa	CGTCAAGTTT	GCTGAGCTGG	TG-----	-ACCAAGATC	GACAGACGA-	-----TC	CGGCAAGGAG	ATG-----	-----GAGAA	GGAGCCCAAG	TTCTTCAAGA	[1150]
033_EF1a_Cyanidioschyzon_merolae	CTGTGCGTTT	GCGGAGCTCT	TA-----	-GAGAAAATC	GATCGCCGC-	-----AC	CGGCAAGAAG	ATC-----	-----GAAGA	GAACCCGAA	AAGGTGAAG	[1162]
034_EF1a_Porphyrha_yezoensis	GTGCAAGTTC	TTCGAGCTGA	TC-----	-CTGAAGATG	GACCGCCGC-	-----TC	GGGCAAGAAG	CTG-----	-----GAGGA	CAGCCCCAAG	ATGATCAAGT	[1162]
035_EF1a_Plasmodium_falciparum	ATGTAATATT	TTAAACATTG	AT-----	-CTGAAGATC	GACAAAACGT-	-----TC	AGGTAAAGTT	GTT-----	-----GAAGA	AAATCCAAA	GCTATCAAGT	[1144]
036_EF1a_Cryptosporidium_parvum	CTGCAAAATC	CAGACTATCA	CT-----	-GCTAAGATG	GACAGAAGA-	-----TC	TGGTAAGGTT	TTG-----	-----GAAGA	AAACCCCAAG	CTTATCAAGT	[1138]
037_EF1a_Tetrahymena_pyriformis	TTGCAAGTTC	GAACACTATCC	AC-----	-GACAAGATC	GACAGAAGA-	-----AC	TGGTAAGTCT	CAA-----	-----GAAGA	AAACCCCAAG	TTATCAAGA	[1153]
038_EF1a_Stentor_coeruleus	TTGCAAGTTC	AACATCATCC	AA-----	-AAGAAGATT	GATAGAAGA-	-----AC	TGTGTCTTCA	TTG-----	-----GAAGA	TGAGCTTAA	GTTGTGAAGA	[1105]
039_EF1a_Phytophthora_parasitica	CTGCAAGTTC	AAGAGATCA	CG-----	-GAGAAGATG	GACCGTCGT-	-----TC	GGGCAAGGTTG	CTC-----	-----GAGAC	TGCCCCCAAG	TTGTCAGAT	[1150]
040_EF1a_Mallomonas_rasilis	CACTAAATTT	AAGAATATTG	AT-----	-GAAAAGATG	GATCGTCGT-	-----AC	TGGAAAGAGT	ATG-----	-----GAGGA	CAACCCGAA	TTGTCAGAA	[1111]
041_EF1a_Phaeodactylum_tricornutum	TTGCAAAATC	GCCCTTATCA	AG-----	-GAGAAGATG	GATCGCCGA-	-----AG	TGGAAAGGTT	TTG-----	-----GAACA	GAACCCCAAG	TTTGTAAAA	[1150]
042_EF1a_Thalassiosira_pseudonana	ATGCCAAATC	ACCGAACCTA	AG-----	-CAGAGAATG	GATCGTACG-	-----AA	CGGCAAGATG	CTT-----	-----GAAGA	CAACCCCTAG	TTGTCAGAG	[1150]
043_EF1a_Malawimonas_jakobiformis	CTGCAAGTTC	AACGAGATCG	TT-----	-GCCAAGGTT	GACCGCCGC-	-----AC	AGGCAAGGCC	AGC-----	-----GAGGA	GGAGCCCAAG	TTATCAAGT	[1081]
044_EF1a_Malawimonas_californiana	GTGCAAGTTC	GCCGAGCTCG	TG-----	-CAGAAGGTC	GACCGCCGC-	-----AC	AGGCAAGGTTG	CTC-----	-----GAGGA	GGAGCCCAAG	TTTGTCAAGT	[1150]
045_EF1a_Trimastix_pyriformis	CTGCAAGTTC	GAGGAGCTCG	TG-----	-ACCAAGATC	GACCGTCGT-	-----AC	TGGCCGCGCC	ACT-----	-----GAGGA	GTCCCCCAAG	ATGGTCAAGA	[1108]
046_EF1a_Dinenympha_exilis	ATGCAAGTTC	AAGGAGATTG	AG-----	-TCAAAGATA	GATCGCCGA-	-----AC	AAACAAGGTTG	CAG-----	-----GAGGA	GAACCCGAA	TTTATAAGA	[1090]
047_EF1a_Monocercomonoides_sp.	ATGCAAGTTC	AAGGAGATCA	CA-----	-GCTAAATTT	GATCGTCGT-	-----AC	CAACAAGATC	CAA-----	-----GAAGA	GAACCCCAAG	TTATCAAGA	[1150]
048_EF1a_Streblomastix_strix	ATGCAAAATTT	AAGGAAATTA	CA-----	-AGCAAAAT	GATAGACGT-	-----AC	AAATAAGAA	TAA-----	-----GAAGC	TAATCTTAA	TCTATTAA	[1099]
049_EF1a_Histomonas_meleagridis	TTGCAAGTTC	GACAAGCTCG	TC-----	-CAACGTATC	GACCGCCGT-	-----CA	CGGTAAAGAA	GTG-----	-----ACAGA	GAACCCGAA	TGGATCCAA	[1150]
050_EF1a_Trichomonas_vaginalis	CTGCAAGTTC	GACAAGCTCA	TC-----	-CAGCGTATC	GATCGTCGC-	-----CA	CGGCAAGAAG	GCT-----	-----ACAGA	GAACCCGAA	TACATTCAGA	[1150]
051_EF1a_Hexamita_inflata	CTGSAAGTTC	GACAAAGTTT	TS-----	-GCCAAGCTC	AACCTCAGA-	-----AC	TTTCAAGGTT	GAA-----	-----ATCGA	GAACCCCAAG	GAGGCCGTC	[1102]
052_EF1a_Spiroplasma_vortens	CTGCGAGTTC	ACCAAGTTCC	TC-----	-CAGAAGCTC	AACCTCCGT-	-----AC	CCTCAAGGCG	GAG-----	-----CAGGA	GAATCCGAC	GAGGCTCGCT	[1105]
053_EF1a_Andalucia_incarcerata	CTGCAAGTTC	GCTGAGCTCA	CC-----	-AAGAAGATC	GATCGCCGA-	-----TC	TGGTAAGGAA	CTC-----	-----GAGGC	TACCCCAAG	TTGTCAGAT	[1099]
054_EF1a_Seculamonas_ecuadoriensis	CTGCGCGTTC	GCCGAGCTCG	TC-----	-CAGAAGATC	GACCGCCGT-	-----AC	CGGTAAAGGTTG	ATG-----	-----GAGGA	GAACCCCAAG	GCCATCAAGT	[1150]
055_EF1a_Reclinomonas_american	GTGCGCGTTC	AACGAGTTG	TG-----	-CAGAAGATC	GATCGCCGT-	-----AC	CGGCAAGGAG	CTT-----	-----GAGGC	CACCCCAAG	TTGTCAGAT	[1150]
056_EF1a_Histiona_aroides	TTGCGCGTTC	GCTGAGTTG	TC-----	-CAGAAGATT	GATCGACGT-	-----AC	CGGTAAAGAA	CTC-----	-----GAAGC	TTCTCCCAAG	TTTGTGAAGT	[1150]
057_EF1a_Jakoba_libera	CTGCGCGTTC	GCCGAGCTGA	TC-----	-ACCAAGATC	GATCGCCGT-	-----AC	CGGTAAAGGAG	CTG-----	-----GAGGC	CTCCCCCAAG	TTGTCAGAT	[1150]
058_EF1a_Stachyamoeba_lipophora	TTGCAAGTTC	GAGGCCATCA	CT-----	-GAGAAGCTC	GACAGAAGG-	-----AC	CGGTAAAGGCT	GTT-----	-----GAGGA	GAACCCCGCC	TTGTCAGAA	[1147]
059_EF1a_Naegleria_gruberi	TTGCAAGTTT	GAAACTATCA	TT-----	-GAAAAGATT	GACAGAAGA-	-----AC	TGGTAAGACT	CAA-----	-----GAAGC	TAACCCGAA	AAGATCAAGA	[1102]
060_EF1a_Naegleria_andersoni	TTGTAAGTTC	GACAAGTTGA	TC-----	-GACAAGATT	GACAGAAGA-	-----AC	TGGTAAGTCC	CAA-----	-----GAAGC	TACCCGAA	AAGATCAAGA	[1120]
061_EF1a_Acrasis_rosea	CTGCAAGTTC	CAAGAGATT	TG-----	-ACCAAGATC	GACAAAACGT-	-----AC	TGGTCAAGAG	ATT-----	-----GAAAA	GCTACCAAG	AGCTTGAAGT	[1099]
062_EF1a_Sawyeria_marylandensis	ATGTAATTTT	GAGAATTTAA	TT-----	-GCATTAGTT	GATAAACGT-	-----AA	TGGTCAAGAA	ATG-----	-----GAAAA	ATCTCCAGT	TCAATTAAGA	[1150]
063_EF1a_Bodo_saltans	CTGCAAGTTC	GCCATGATCG	AG-----	-TCCAAGATT	GACCGTCGT-	-----TC	CGGCAAGGAG	GTC-----	-----GAGAA	GGAGCCCAAG	GCGATCAAGT	[1090]
064_EF1a_Rhynchobodo_sp.	CTGCAAGTTC	CAGGAGATCG	AG-----	-TCCAAGATT	GACCGTCGT-	-----TC	TGGTAAGGAG	CTC-----	-----GAGAA	GGAGCCCAAG	GCCATCAAGA	[1084]
065_EF1a_Rhynchomonas_nasuta	GTGCAAGTTC	GCCGAGATCG	AG-----	-TCCAAGATT	GACCGTCGT-	-----TC	GGGCAAGGTC	ATG-----	-----GAGGA	GGAGCCCAAG	TTATCAAGA	[1147]
066_EF1a_Dimastigella_trypaniformis	GTGCAAGTTC	ACGAGCTCG	AG-----	-GCCAAGATT	GACCGCCGT-	-----AC	CGGCAAGGAA	TCC-----	-----GAGGC	GTCCCCCAAG	TTGTCAGAA	[1105]

	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	
067_EF1a_Perkinsiella_amoebae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[915]
068_EF1a_Leishmania_major	GTGCCGCTTC	GC GGAAATCG	AG-----	-TCCAAGATC	GACCGCCGC-	-----TC	CGGCAAGGAG	CTG-----	-----GAGAA	GAACCCCAAG	GC GATCAAGT	[1150]
069_EF1a_Leishmania_infantum	GTGCCGCTTC	GC GGAAATCG	AG-----	-TCCAAGATC	GACCGCCGC-	-----TC	CGGCAAGGAG	CTG-----	-----GAGAA	GAACCCCAAG	GC GATCAAGT	[1150]
070_EF1a_Leishmania_braziliensis	GTGCCGCTTC	GC GGACATCG	AG-----	-TCCAAGATC	GACCGCCGC-	-----TC	CGGCAAGGAG	CTG-----	-----GAGAA	GAACCCCAAG	GC GATCAAGT	[1150]
071_EF1a_Leishmania_tarentolae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[1032]
072_EF1a "L. tarentolae" FJ807257	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[915]
073_EF1a_Leptomonas_costaricensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[594]
074_EF1a_Leptomonas_bifurcata	AT-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[832]
075_EF1a_Leptomonas_podlipaevi	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[814]
076_EF1a_Herpetomonas_muscarum	GTGCAAGTTC	GCCGACATCG	AG-----	-TCCAAGATC	GACCGCCGC-	-----TC	CGGTAAGGTG	CTC-----	-----GAGGA	GCACCCCAAG	TTGTCGTAAGT	[1063]
077_EF1a_Herpetomonas_pessoai	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[771]
078_EF1a_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[766]
079_EF1a_Trypanosoma_brucei	CTGCAAGTTC	GCGGAGATCG	AG-----	-TCGAAGATC	GACCGTCGC-	-----TC	TGGCAAGGAG	CTG-----	-----GAGAA	GGCTCCCAAG	TCGATCAAGT	[1150]
080_EF1a_T. brucei_equiperdum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[774]
081_EF1a_Trypanosoma_aviium	GTGCAAGTTC	GCTGAGATCG	AG-----	-TCCAAGATT	GACCGTCGC-	-----TC	TGGTAAGGAG	CTG-----	-----GAGAA	GAACCCCAAG	GCCATCAAGT	[1096]
082_EF1a_Trypanosoma_cruzi	GTGCAAGTTC	GCCGAGATCG	AG-----	-TCCAAGATC	GACCGCCGC-	-----TC	GGGCAAGGAG	CTT-----	-----GAGAA	GAACCCCAAG	TCGATCAAGT	[1150]
083_EF1a_Trypanosoma_rangeli	GTGCAAGTTT	GCCGAGATCG	AG-----	-TCCAAGATC	GACCGCCGC-	-----TC	TGGCAAGGAG	CTT-----	-----GAGAA	GAACCCCAAG	TCGATCAAGT	[1150]
084_EF1a_Ploetia_costatum	GTGCAAGTTC	GCCACCATTC	AA-----	-AACAAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TCCATCAAGT	[1147]
085_EF1a_Ploetia_oblonga_cDNA	GTGCAAGTTC	GCCACCATTC	AA-----	-AACAAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TCCATCAAGT	[1147]
086_EF1a_Ploetia_oblonga_gDNA	GTGCAAGTTC	GCCGAAATCA	AG-----	-ACAAAGATT	GACCGTCGT-	-----TC	CGGCAAGGAG	CTC-----	-----GAGGT	TGAGCCCAAG	GCCATCAAGA	[1147]
087_EF1a_Entosiphon_sp.	CTGCAAGTTC	GCTGAGATCT	TG-----	-ACCAAGATC	GACCGCCGC-	-----TC	TGGCAAGGAG	CTC-----	-----GAGCA	GGAGCCCAAG	GCTATCAAGT	[1147]
088_EF1a_Entosiphon_sulcatum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[915]
089_EF1a_Peranema_tricophorum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[938]
090_EF1a_P. tricophorum_FJ807243	TTGCAAAATTC	TACGAGATCA	AG-----	-ACCAAGATC	GACAGACGT-	-----AG	TGGAAAGGAA	TTG-----	-----GAAGC	CGAACCAAAA	TTTATCAAGA	[1141]
091_EF1a_Colacium_elongatum	TTGCAAGTTC	GCCACCATTC	AG-----	-ACCAAGATT	GACCGACGA-	-----TC	TGGAAAGGAG	TTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1147]
092_EF1a_Colacium_mucronatum	TTGTAAGTTC	GCCACCATTC	AG-----	-ACCAAGATT	GACCGTCGC-	-----TC	TGGCAAGGAG	TTG-----	-----GAGGC	CGAGCCCAAG	TTTATCAAGT	[1093]
093_EF1a_Colacium_sp.	CTGCAAGTTC	GCCACCATTC	AG-----	-ACCAAGATT	GACCGTCGT-	-----TC	TGGTAAGGAG	CTG-----	-----GAGGC	CGAGCCCAAG	TTTATCAAGT	[1147]
094_EF1a_Colacium_vesiculosum	TTGTAAGTTC	GCCACCATTC	AG-----	-ACCAAGATT	GACCGTCGC-	-----TC	TGGCAAGGAG	TTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1105]
095_EF1a_Cryptoglena_pigra	TTGCAAGTTT	GCCACCATTC	AG-----	-ACAAAGATC	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
096_EF1a_Cryptoglena_skujai	TTGCAAGTTC	GCCACCATTC	AG-----	-ACCAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
097_EF1a_Discoplistis_adunca	ATGCAAGTTT	GCAACCATTC	AG-----	-ACCAAGATT	GACCGACGT-	-----TC	CGGCAAGGAG	CTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1147]
098_EF1a_Euglena_clara	TTGCAAGTTT	GCAACCATTC	AA-----	-ACCAAAATC	GATCGCCGA-	-----TC	TGGCAAGGAG	TTG-----	-----GAGGC	TGAACCCAAAG	TTTATCAAGT	[1027]
099_EF1a_Euglena_deses	TTGCAAGTTC	GCGACAATTC	AG-----	-ACCAAGATT	GACCGACGT-	-----TC	TGGCAAGGAG	TTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1147]
100_EF1a_Euglena_exilis	CTGCAAAATTC	GCCACCATTC	AG-----	-ACCAAGATT	GATCGTCGA-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	TGAACCCAAAG	TTTATCAAGT	[1027]
101_EF1a_Euglena_geniculata	GTGCAAGTTC	GCCACCATTC	AA-----	-ACCAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
102_EF1a_Euglena_gracilis	CTGCAAGTTC	GCCACCATTC	TG-----	-AACAAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	CGAGCCCAAG	TTTATCAAGT	[1147]
103_EF1a_E. gracilis_X16890	GTGCAAGTTC	GCAACCATTC	AA-----	-ACAAAGATT	GACAGGCCG-	-----TC	TGGAAAGGAA	TTG-----	-----GAGGC	AGAGCCCAAG	TTTATCAAGT	[1150]
104_EF1a_E. gracilis_FJ807238	CTGCAAGTTC	GCCACCATTC	TG-----	-AACAAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	CGAGCCCAAG	TTTATCAAGT	[1150]
105_EF1a_Euglena_granulata	CTGCAAGTTT	GCAACCATTC	AG-----	-ACAAAGATC	GATCGCCGA-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	TGAACCCAAAG	TTTATCAAGT	[1147]
106_EF1a_Euglena_hiemalis	TTGCAAGTTC	GCCACCATTC	AG-----	-ACCAAGATT	GACCGACGT-	-----TC	TGGGAAGGAG	CTG-----	-----GAGGC	AGAGCCCAAG	TTTATCAAGT	[1147]
107_EF1a_Euglena_longa	CTGCAAGTTC	GCAACCATTC	AG-----	-ACCAAGATT	GACCGACGT-	-----TC	CGGAAAGGAG	CTG-----	-----GAGGC	AGAGCCCAAG	TTTATCAAGT	[664]
108_EF1a_Euglena_stellata	CTGCAAGTTT	GCCACAATTC	AG-----	-ACCAAAATT	GATCGCCGT-	-----TC	TGGTAAGGAG	CTC-----	-----GAGGC	TGAGCCCAAG	TTTATTAAT	[1027]
109_EF1a_Euglena_viridis	YTGCAAGTTC	GCTACCATTC	AG-----	-ACCAAGATT	GACCGCCGC-	-----TC	TGGCAAGGAA	TTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1105]
110_EF1a_Eutreptia_pertyi	CTGCAAGTTC	GCAACCATTC	AG-----	-ACAAAGATC	GATCGCCGA-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	TGAACCCAAAG	TTTATCAAGT	[1147]
111_EF1a_Eutreptiella_pomquetensis	GTGCAAGTTC	GCCACCATTC	AA-----	-AACAAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[757]
112_EF1a_Lepocinclis_acus_var._major	TTGCAAGTTT	GCCACCATTC	AG-----	-ACCAAGATT	GACCGTCGA-	-----TC	TGGTAAGGAG	CTG-----	-----GAGGC	AGAGCCCAAG	CACATCAAGT	[1147]
113_EF1a_Lepocinclis_ovum	TTGCAAGTTC	GCAACAATTC	AG-----	-ACCAAGATT	GACCGACGA-	-----TC	TGGTAAGGAA	TTG-----	-----GAGGC	TGAACCCAAAG	TTTATCAAGT	[1027]
114_EF1a_Lepocinclis_spyrogyroides_clone_38	GTGCAAGTTT	GCGACCATTC	AG-----	-ACCAAGATT	GATCGCCGT-	-----TC	TGGCAAGGAA	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[460]
115_EF1a_L. spyrogyroides_clones_1_&_4	TTGCAAGTTC	GCCACCATTC	AG-----	-ACCAAGATT	GACCGCCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
116_EF1a_Lepocinclis_steinii	CTGCAAGTTT	GCCACCATTC	AG-----	-AACAAAGATT	GATCGCCGT-	-----TC	CGGCAAGGAG	CTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1027]
117_EF1a_Lepocinclis_triapteris	CTGCAAGTTC	GCCACCATTC	AG-----	-ACAAAGATT	GACCGACGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	CGAGCCCAAG	TTTATCAAGT	[1147]
118_EF1a_Monomorphina_aenigmatica	ATGCAAGTTC	GCCACTATTC	AG-----	-ACCAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	TTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1147]
119_EF1a_Monomorphina_pyrum	TTGCAAGTTC	GCCACCATTC	AG-----	-ACGAAGATT	GACCGCCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
120_EF1a_Monomorphina_sp.	CTGCAAGTTC	GCCACCATTC	AG-----	-ACGAAATTC	GACCGACGT-	-----TC	TGGCAAGGAG	CTC-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
121_EF1a_Phacus_orbicularis	ATGCAAGTTT	GCTGAGATTC	AG-----	-ACCAAGATT	GATCGCCGT-	-----TC	TGGGAAGGAG	TTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1105]
122_EF1a_Phacus_similis	CTGCAAGTTC	GCCACCATTC	AG-----	-ACCAAGATT	GACCGCCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
123_EF1a_Strombomonas_acuminata	TTGCAAGTTT	GCAACCATTC	AG-----	-ACCAAGATT	GACCGCCGT-	-----TC	AGGCAAGGAG	CTT-----	-----GAGGC	GGAAACCAAG	TTTATCAAGT	[1150]
124_EF1a_Strombomonas_borysteniensis	CTGCAAGTTT	GCCACCATTC	AG-----	-ACCAAGATT	GACCGCCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1027]
125_EF1a_Strombomonas_napiformis	CTGCAAGTTT	GCCACCATTC	AG-----	-ACGAAGATT	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
126_EF1a_Strombomonas_ovalis	CTGCAAGTTT	GCCACCATTC	AG-----	-ACCAAGATT	GACCGCCGT-	-----TC	CGGCAAGGAG	CTG-----	-----GAGGC	GGAGCCCAAG	TTTATCAAGT	[1147]
127_EF1a_Strombomonas_triquetra	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[891]
128_EF1a_Trachelomonas_conradi	CTGCAAGTTC	GCAACCATTC	AG-----	-ACCAAGATT	GACCGACGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	CGAACCCTAAG	TTTATCAAGT	[1147]
129_EF1a_Trachelomonas_echinata	CTGCAAGTTT	GCAACCATTC	AG-----	-ACCAAGATC	GACCGTCGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	AGAACCCTAAG	TTTATCAAGT	[1147]
130_EF1a_Trachelomonas_ilefevrei	CTGCAAGTTC	GCAACAATTC	AG-----	-ACAAAGATT	GACCGACGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1147]
131_EF1a_Trachelomonas_plancetonica	CTGCAAGTTC	GCAACAATTC	AA-----	-ACAAAGATT	GACCGACGT-	-----TC	TGGCAAGGAG	CTG-----	-----GAGGC	TGAGCCCAAG	TTTATCAAGT	[1027]
132_EF1a_Trachelomonas_similis	CTGCAAGTTT	GCAACCATTC	AG-----	-ACAAAGATT	GATCGTCGT-	-----TC	AGGCAAGGAR	TTG-----	-----GAAGC	TGAGCCCAAG	TTTATCAAGT	[1027]

	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	
133_EFL1a_Trachelomonas_volvocina	TTGCAAGTTT	GCAACCATCC	AG-----	-ACCAAGATT	GATCGCCGT-	-----TC	TGGCAAGGAG	CTC-----	-----GAGGC	CGAGCCCAAG	TTCATCAAGT	[1147]
134_EFLenvironmental_sample_EU810232	CGTGAAGATG	GTCAAGATCG	ACTGGAAGAT	CGGAAAGGAG	ACCGGAGGCC-	-----	-AAGAAAG--	-----	-----CTCGA	TGACCCCGTG	TCCATCAAGA	[1240]
135_EFLCreolimax_fragrantissima	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[994]
136_EFLSphaeroforma_arctica	CGTGAAGATG	ACCTTCATTG	ACTGGAAGGT	CGGTAAAGGAG	ACCGGTGGC-	-----	-ACCAAG--	-----	-----GCCGA	GAACCCCGCT	GGTCTTAAAG	[1240]
137_EFLMonosiga_brevicollis	GTGCCGCATG	ACCGCCATCA	ACTGGAAGAT	TGGCAAGGAG	ACCGGTGGC-	-----	-AAGAAAG--	-----	-----GCCGA	GGACCCCGTT	GGCCTCAAGA	[1240]
138_EFLXenopus_laevis	CGTCAAGCTG	ACCAAGATCG	CCTGGAAGGT	TGGCAAGGAG	ACTGGTGGC-	-----	-TCCAAG--	-----	-----ATGGA	GAACCCCGAG	AGCATCAAGG	[1240]
139_EFLSpizellomyces_punctatus	TTGCCGTATG	AAGGAGATCA	AGTGGAAAGT	GGGTAAAGGAG	ACTGGTATG-	-----	-AAGAAAG--	-----	-----CTTGA	GGGTGCCACC	TCCTCAAGG	[1246]
140_EFLGaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
141_EFLAllomyces_macrogynus	CTGCAAGATC	ACCAAGATCA	ACTGGAAGAT	TGGCAAGGAG	ACCGGTCGC-	-----	-AAAAAG--	-----	-----ATGGA	GGGCCCGCAC	TCGCTCAAGG	[1246]
142_EFLBlastocladiella_emersonii	CTGCAAGATC	ACCAAGATCA	ACTGGAAGAT	CGGCAAGGAG	ACCGGTCGC-	-----	-AAGAAAG--	-----	-----ATGGA	GTCCCCCAC	TCGCTCAAGG	[1246]
143_EFLCatenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
144_EFLCoelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[506]
145_EFLEntomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
146_EFLBasidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
147_EFLConidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[503]
148_EFLTetraselmis_tetrathele	CTGCAGGATC	ATGAAGCTGG	TCAACAAGAA	GGGCAAGGAG	ACTGGCGGC-	-----	-AAGATG--	-----	-----CTGGA	GGACCCCCAC	TCCTCAAGT	[1819]
149_EFLMesostigma_viride	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[879]
150_EFLOstreococcus_tauri	GTGCAAGCTC	ACGGCGCTCA	ACTTCAAGGT	TGGCAAGGAA	ACCGGAGGT-	-----	-AAGAAAG--	-----	-----CTCGA	GTCCCCGATG	TCGCTGAAGT	[1249]
151_EFLNephroselmis_olivacea	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[957]
152_EFLMicromonas_pusilla	GTGCAAGATG	ACCGAGCTCG	TCTTCAAGGT	CGGCAAGGAG	ACCGGCGGC-	-----	-AAGAAAG--	-----	-----CTGGA	GTCCCCCATG	GCCTCAAGT	[1198]
153_EFLParachlorella_kesslerii	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[942]
154_EFLHelicosporidium_sp.	CTGCCGCATG	ACCAAGATCA	ACTGGAAGGT	CGGCAAGGAG	ACCGGTGGC-	-----	-AAGAAAG--	-----	-----CTGGA	GGACCCCCAC	TCCTCAAGG	[1261]
155_EFLGonium_pectorale	CTGCCGCATG	CAGGCGATCA	ACTGGAAGGT	CGGCAAGGAG	ACCGGCGGC-	-----	-AAGAAAG--	-----	-----ATGGA	GGGCCCGGTC	GGCCTGAAGG	[1243]
156_EFLChlamydomonas_reinhardtii	TTGCCGCATC	AGCGGCATCA	ACTGGAAGGT	CGGCAAGGAG	ACCGGCGGC-	-----	-AAGAAAG--	-----	-----ATGGA	GAACCCCGTC	GGCCTGAAGG	[1243]
157_EFLScenedesmus_obliquus	CTGCAAGATC	TCCAAGATCG	ACTGGAAGGT	GGGCAAGGAG	ACTGGTGGC-	-----	-AAGAAAG--	-----	-----CTGGA	GGCCCCCAC	GCCTCAAGG	[1243]
158_EFL"Oryza_sativa"	CTGCAAGATC	ACAGAGATCA	AGTGGAAAGT	CGGCAAGGAG	ACTGGTAAAC	-----	-AAGAAAG--	-----	-----CTGGA	GGAGCCCCAC	GGCCTGAAGG	[1243]
159_EFLChlorococcum_sp.	TTGCCGTGTG	ACCAACCTGG	ACTGGAAGGT	GGGCAAGGAG	ACTGGCGGC-	-----	-AAGAAAG--	-----	-----CTCGA	GGAGCCCCAC	TCCTCAAGT	[1192]
160_EFLOchlochaete_hystrix	CTGCAGGATC	TCCGGCATCA	ACTGGAAGGT	GGGCAAGGAG	ACCGGTGGC-	-----	-AAGAAAG--	-----	-----CTGGA	GGCCCCCAC	AGTCTGAAGG	[1201]
161_EFLUlva_intestinalis	CTGCCGTGTC	ACCAAGATTG	ACTGGAAGGT	CGGCAAGGAG	ACTGGTGGC-	-----	-AAGAAAG--	-----	-----CTCGA	GGAGCCCCAC	TCGCTGAAGT	[1192]
162_EFLUrospora_sp.	GTGCCGCGTG	TCGGCGATCG	ACTGGAAGGT	CGGCAAGGAG	ACCGGTGGC-	-----	-AAGAAAG--	-----	-----CTCGA	CGAGCCGCAC	AGCCTCAAGT	[1045]
163_EFLGymnochlorella_stellata_ppc	ATGCAAGCTG	AAGAAGATAA	ATTGGAAGAT	GGGTAAAGGAG	ACTGGTGGC-	-----	-AAGAAAG--	-----	-----GTTGA	GAATCCCGCC	ATGCTTAAAG	[1279]
164_EFLLotharella_amoebiformis_ppc	TTGCAAGCTG	AAGAAGATAA	ACTGGAAGAT	GGGCAAGGAG	ACAGGAGGAA	-----	-AAGAAAG--	-----	-----CTAGA	GAACCCCAAG	CTACTCAAGC	[1024]
165_EFLChlorarachnion_reptans_ppc	CTGCAGAGTT	AAAAAGATAA	ACTGGAAGAT	GGGCAAGGAG	ACCGGAGGG-	-----	-AAGAAAG--	-----	-----CTCGA	AAATCCGCCG	CATCTGAAGC	[1051]
166_EFLChondrus_crispus	CTGCCGAATC	AAATGCATCA	ACTGGAAGAT	GGGCAAGGAG	ACTGGAAGT-	-----	-GCCAAG--	-----	-----GTCGA	GAACCCCGAT	TACATCAAGG	[1240]
167_EFLGracilaria_changii	CTGCAAGATC	AAGTCATCA	ACTGGAAGAT	GGGCAAGGAG	ACTGGTGGT-	-----	-GCCAAG--	-----	-----GTCGA	GCACCCAGAA	TACATTAAGG	[1240]
168_EFLCryptomonas_ovata	GTGCCGAGATG	TCGGTGATCG	ACTGGAAGAT	GGGCAAGGAG	ACTGGTGGC-	-----	-GCCAAG--	-----	-----GCGBA	GAGCCCGCCC	CACCTCAAGG	[1201]
169_EFLGuillardia_theta	TTGCCGTATG	ACCGTATCG	ACTGGAAGAT	GGGTAAAGGAG	ACCGGAGGA-	-----	-CAGAAG--	-----	-----CTCGA	GAACCCCTCT	CATCTCAAGG	[1246]
170_EFLRhodomonas_salina	ATGCAGGATG	TCTGTCATCG	ACTGGAAGAT	GGGTAAAGGAG	ACCGGTGGC-	-----	-CAGAAG--	-----	-----CTCGA	GAGCCCGCCC	CATTGAAGG	[1243]
171_EFLPavlova_lutheri	GTGCCGCATG	ACGGAGATCA	AGTTCAAAGT	CGGCAAGGAG	ACG6GG--	-----	-----	-----	-----	-----	-----	[1206]
172_EFLChrysochromulina_sp.	TTGCCGCGCG	ACGGGCTTCA	ACTTCAAAGT	CGGCAAG--	-----	-----	-----	-----	-----	-----	-----	[1017]
173_EFLIsochrysis_galbana	ATGCCGCGCG	ACCGCTTTTG	TTTTCAAAGT	CGGCAAGGAG	ACTGGCGGC-	-----	-AAGAAAG--	-----	-----ATGGA	GGGCCCCAC	TCGCTCAAGT	[1243]
174_EFLEmiliana_huxleyi	GTGCCGCGCG	ACCGCTTCA	AGTTCAAAGT	CGGCAAGGAG	ACCGGCGGC-	-----	-AAGAAAG--	-----	-----ATGGA	GTGCGGCAC	TCGCTCAAGT	[1249]
175_EFLAchnanthes_kuwaitensis	GTGCAAGATG	ACCAAGATCT	TGTGGAAGCA	GTCCAAGAG	ACTGGAAGT-	-----	-GCCAAG--	-----	-----GTTGA	GAACCCCGCC	GAGCTTTCCC	[1150]
176_EFLAsterionella_glacialis	ATGCAAGATG	ACCAAGATCC	TTTGGAAAGCA	ATCCAAGAG	ACTGGTGA-	-----	-GAAAAG--	-----	-----GTTGA	AAACCCCGCC	GAACTTTCTC	[1150]
177_EFLDetonula_confervacea	ATGCAAGATG	AGCAAGATCC	TGTGGAAGAT	GGGCAAGAG	ACCGGTGA-	-----	-CAGAAG--	-----	-----CTCGA	CAACCCCATC	GAACTTTCTC	[1150]
178_EFLThalassiosira_pseudonana	CTGTAAAGATG	ACCAAGATCT	TGTGGAAGAT	GGGAAAGAG	ACTGGAGGA-	-----	-CAGAAG--	-----	-----GTCGA	CAACCTCTCT	GAGCTTTCCG	[1213]
179_EFLPerkinsus_marinus	TTGTGCTTTG	GCTGGCATTA	AATGGAAGGT	TGGCAAGGAG	ACTGGTGGC-	-----	-AAGAAA--	-----	-----ATGGA	GGACCTCAT	GCTCTCAAGT	[1258]
180_EFLOxyrrhis_marina	GTGCCGAATT	TCTGGGTTGA	ACTGGAAGGT	TGGCAAGGAG	ACCGGAGGC-	-----	-AAGAAAG--	-----	-----ATGCC	TGAGCCCAT	TCTCTGAAGT	[1252]
181_EFLKarlodinium_micrum	TTGTGCGATC	TCCAAGCTCA	AATGGAAGAT	GGGTAAAGAG	ACCGGCGGC-	-----	-AAGAAAG--	-----	-----ATGGA	AGATCCTCAT	GCATTGAAAT	[1246]
182_EFLAmphidinium_carterae	TTGCCGTATC	TACGCTTTGA	AGTGGAAAGT	GGGCAAGGAG	ACAGGCGGC-	-----	-AAGAAA--	-----	-----ATGGA	GGACCCCGAC	TCCTTGAAGT	[1246]
183_EFLHeterocapsa_triquetra	TTGCCGTATC	TCCGCCCTGA	AGTGGAAAGT	GGGCAAGGAG	ACTGGTGGC-	-----	-AAGAAAG--	-----	-----ATGGA	GGATCCCCAC	TCGCTGAAGT	[1246]
184_EFLPlanoglabratella_operculis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[961]
185_EFLReticulomyxa_filosa	TTGCCAAATG	TTGGAATCA	AATGGAAGAT	GGGAAAGCA	ACAAACAAT-	-----	-GCAAAA--	-----	-----GTTGA	GGGCGCACT	TTTATTGAAG	[1186]
186_EFLLotharella_amoebiformis	TTGCAAGCTC	TCCAAAATTG	TCTGGAAGGT	GACCAAGGCC	AACCTTAAGC	AGTGAAGAA	AAAGACTGAT	ATGGAAAAC	ACAAAGAGAT	GAACCCCAAA	TTTATTGAAG	[1351]
187_EFLGymnochlorella_stellata	TTGCCGTGTC	AGCAAGATTC	TCTGGAAGGT	TACCAAGGCT	AACCTGAAGA	CCATCAAGAA	AAAGGCTGAG	ATGGAGAAT	TCAAGGAGAA	GGATGCAAAA	TTTGTTCAGA	[1324]
188_EFLChlorarachnion_reptans	TTGCAAAATG	ACCAAAATTC	ACTGGAAGGT	CACCAAGACT	AACCAAAAGC	AAATGAAGTC	CAAGAAAGAT	CTTGAGAACT	TCAAGGAGGC	AGATGCTAAA	TTTGTTCAGA	[1255]
189_EFLDiplonema_ambulator	CGTGAAGATG	ATGGAGCTTG	TCTGGAAGAT	GGGCAAGGAG	ACC6GGGGT-	-----	-AAGAAAG--	-----	-----GCCGA	GGCCCCCAC	TGCTTGAAGG	[844]
190_EFLDiplonema_papillatum	TGTGAAGATG	TCGCAGCTTG	TCTGGAAGAT	GGGCAAGGAG	ACC6GGCGA-	-----	-AAGAAAG--	-----	-----GCTGA	GGAACTCTAC	TGCTTGAAGG	[1216]
191_EFLRhynchopus_euleeides	CGTGGCGATG	ACCCAGCTGG	TGTGGAAGAT	GGGCAAGGAG	ACC6GGCGG-	-----	-AAGAAAG--	-----	-----GCCGA	GGAGCCCAAC	TGCTTGAAGG	[859]
192_EFLPetalomonas_cantuscygni	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[924]
193_EFLTrypanoplasma_borrelii	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[1017]
194_EFLNeobodo_salien	TTGCCGCTG	TCCGAGATTC	GCTTCAAGGC	GGGCAAGAG	ACC6GGCGC-	-----	-CAGAAG--	-----	-----GTCGA	GTGCGCCGAG	TTCATCGAGA	[1213]

	1550	1560	1570	1580	1590	1600	1610	1620	1630	1640	1650	
067_EF1a_Perkinsiella_amoebae	---	---	---	---	---	---	---	---	---	---	---	[915]
068_EF1a_Leishmania_major	CTGGCGATGC	C6CGATC---	GTGAAGATGG	TGCCGCAGAA	GCCGATGTGC	GTGGAGGTGT	TCAACGACTA	CGCGCCGCTG	GGCCGCTTTG	CCGTGCGCGA	CATGCGCCAG	[1257]
069_EF1a_Leishmania_infantum	CTGGCGATGC	C6CGATC---	GTGAAGATGG	TGCCGCAGAA	GCCGATGTGC	GTGGAGGTGT	TCAACGACTA	CGCGCCGCTG	GGCCGCTTTG	CCGTGCGCGA	CATGCGCCAG	[1257]
070_EF1a_Leishmania_braziliensis	CTGGCGACGC	C6CCATC---	GTGAAGATGG	TGCCGCAGAA	GCCGATGTGC	GTGGAGGTGT	TCAACGACTA	CCCGCCGCTG	GGCCGCTTCG	CCGTGCGCGA	CATGCGCCAG	[1257]
071_EF1a_Leishmania_tarentolae	---	---	---	---	---	---	---	---	---	---	---	[1032]
072_EF1a "L. tarentolae" FJ807257	---	---	---	---	---	---	---	---	---	---	---	[915]
073_EF1a_Leptomonas_costaricensis	---	---	---	---	---	---	---	---	---	---	---	[594]
074_EF1a_Leptomonas_bifurcata	---	---	---	---	---	---	---	---	---	---	---	[832]
075_EF1a_Leptomonas_podlipaevi	---	---	---	---	---	---	---	---	---	---	---	[814]
076_EF1a_Herpetomonas_muscarum	CCGGCGATGC	C6CCATG---	GTCAAGATGG	TGCCGCAGAA	GCCCATGTGC	GTGGAGGTGT	TCAACGACTA	CCCGCCTCTG	GGCC-----	---	---	[1144]
077_EF1a_Herpetomonas_pessoai	---	---	---	---	---	---	---	---	---	---	---	[771]
078_EF1a_Blastocrithidia_culicis	---	---	---	---	---	---	---	---	---	---	---	[766]
079_EF1a_Trypanosoma_brucei	CTGGCGACGC	C6CGATC---	GTGCGCATGG	TGCCGCAGAA	GCCTATGTGC	GTGGAGGTCT	TCAACGACTA	CGCGCCACTC	GGCCGCTTTG	CCGTGCGTGA	CATGCGCCAG	[1257]
080_EF1a_T. brucei_equiperdum	---	---	---	---	---	---	---	---	---	---	---	[774]
081_EF1a_Trypanosoma_aviium	CAGGTGATGC	CGCTCTG---	GTCCGCGATGG	TGCCACAGAA	GCCCATGTGC	GTGGAGGTGT	TCAACGATTA	CGCGCCCTCT	GGTCGCTTTG	CCGTGCGTGA	C-----	[1194]
082_EF1a_Trypanosoma_cruzi	CCGGTGACGC	TGCCATG---	GTGCGCATGG	TGCCGCAGAA	GCCCATGTGC	GTGGAGGTGT	TCAACGACTA	CGCTCCTCTT	GGCCGCTTTG	CCGTGCGTGA	CATGCGCCAG	[1257]
083_EF1a_Trypanosoma_rangeli	CCGGTGACGC	C6CCCTC---	GTGCGCATGG	TGCCGCAGAA	GCCCATGTGC	GTGGAGGTGT	TCAACGACTA	CGCGCCTCTC	GGCCGCTTTG	CCGTGCGTGA	CATGCGCCAG	[1257]
084_EF1a_Ploetia_costatum	CTGGTGATGC	TGCCATC---	GTCTGTATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TACCCGACTA	TGCCCCCTCT	GGTCC-----	---	---	[1229]
085_EF1a_Ploetia_oblonga_cDNA	CTGGTGATGC	TGCCATC---	GTCTGTATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TACCCGACTA	YCCCCCTCTC	GGCCC-----	---	---	[1229]
086_EF1a_Ploetia_oblonga_gDNA	CGGGTGACGC	CGCTATC---	GTG-----	---	---	---	---	---	---	---	---	[1167]
087_EF1a_Entosiphon_sp.	CCGGTGATGC	C6CCATG---	GTGCTGTATGC	GCCCGTTGAA	GCCGATGTGC	GTGGAGTCGT	TACCCGATTA	CGCCCCCTTA	GGACC-----	---	---	[1229]
088_EF1a_Entosiphon_sulcatum	---	---	---	---	---	---	---	---	---	---	---	[915]
089_EF1a_Peranema_tricophorum	---	---	---	---	---	---	---	---	---	---	---	[938]
090_EF1a_P. tricophorum_FJ807243	GTGGAGACGC	TGCCATC---	GTCTCATGA	AACCCACCAA	GCCTATGTGC	GTGGAGATT	TCAGTGACTA	CCCCCACTC	GGAAAGATTG	CCGTCCGTGA	CATGAGACAA	[1248]
091_EF1a_Colacium_elongatum	CTGGTGATGC	TGCAATT---	GTGCTGTATGA	AGCCCCAGAA	GCCCATGTGC	GTGAGTCCT	TCAGTGACTA	CCCTCCTCTG	GGTCGGTTTG	CTGTCCGTGA	CATGCGCCAA	[1254]
092_EF1a_Colacium_mucronatum	CTGGTGATGC	---	---	---	---	---	---	---	---	---	---	[1103]
093_EF1a_Colacium_sp.	CTGGTGATGC	C6CCATC---	GTCTCATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TYACYGACTA	CCCYCCCC--	---	---	---	[1222]
094_EF1a_Colacium_vesiculosum	CTGGTGATGC	TGCCATT---	GTCTGTATGA	AGCCACAGAA	GCCCATGTGC	GTGGAGTCCT	TCACCGACTA	CCCTCCTCTG	GGGCGCTT--	---	---	[1190]
095_EF1a_Cryptoglena_pigra	CTGGTGATGC	TGCAATT---	GTGCTGTATGA	AGCCCCAGAA	GCCAATGTGC	GTGGAATCCT	TACCCGATTA	TCCCCCTT	GGACC-----	---	---	[1229]
096_EF1a_Cryptoglena_skujai	CTGGCGATGC	C6CCATC---	GTGCTGTATGA	AGCCCCAGAA	GCCGATGTGC	GTGGAATCCT	TACCCGATTA	SCCCCC--	---	---	---	[1221]
097_EF1a_Discoplastis_adunca	CTGGTGATGC	GGCCATT---	GTCTCATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAATCCT	TACCCGACTA	CCCCCCCTCT	GGGCC-----	---	---	[1229]
098_EF1a_Euglena_clara	CAGGCGATGC	TGCTATC---	GTCTCATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TCAGTGATTA	CCCGCCCTCT	GGGCGGTTTG	CCGTGCGTGA	CATGCGCCAA	[1134]
099_EF1a_Euglena_deses	CTGGTGATGC	C6CCATC---	GTCTGTATGA	AGCCGCGAGAA	GCCCATGTGC	GTGGAGTCCT	TCACCGAYTA	CCCCCCCCCT	GGGCC-----	---	---	[1229]
100_EF1a_Euglena_exilis	CAGGAGATGC	TGCAATC---	GTGCTGTATGA	AGCCCCAGAA	GCCAATGTGC	GTGGAATCCT	TCAGTGATTA	TCCACCTCTT	GGACGATTTG	CYGTGCGTGA	CATGAGGCAG	[1134]
101_EF1a_Euglena_geniculata	CTGGTGATGC	TGCCATC---	GTCTGTATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TACCCGACTA	CCCCCCCTTA	GGGCC-----	---	---	[1229]
102_EF1a_Euglena_gracilis	CTGGTGATGC	TGCCATC---	GTATCATGA	AACCCAGAA	GCCCATGTGC	GTGGAGTCGT	TACCCGACTA	SCCCCCCTT	GGACC-----	---	---	[1229]
103_EF1a_E. gracilis_X16890	CAGGTGATGC	TGCCATT---	GTCTGTATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCGT	TCAGTGATTA	CCCCCATTG	GGCGTTTCTT	GT---GGCGA	CATGCGACAA	[1254]
104_EF1a_E. gracilis_FJ807238	CTGGTGATGC	TGCCATC---	GTATCATGA	AACCCAGAA	GCCCATGTGC	GTGGAGTCGT	TCAGTGATTA	CCCCCCCTG	GGACGGTTTG	CCGTGCGTGA	CATGCGGCAG	[1257]
105_EF1a_Euglena_granulata	CTGGTGATGC	TGCCATC---	GTGCTGTATGA	AGCCTCAGAA	GCCATGTGC	GTGGAGTCCT	TACCCGACTA	CCCCCCCTTA	GGACC-----	---	---	[1229]
106_EF1a_Euglena_hiemalis	CTGGTGATGC	TGCCATT---	GTCTGTATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TCACCGACTA	CCCCCCCCCT	GGACC-----	---	---	[1229]
107_EF1a_Euglena_longa	CTGGTGATGC	TGCCATC---	GTCTGTATGA	AGCCACAGAA	GCCCATGTGC	GTGGAATCCT	TCAGTGATTA	CCCCCCACTG	GGACGGTTTG	CTGTTGCGTGA	CATGCGGCAG	[771]
108_EF1a_Euglena_stellata	CTGGAGATGC	TGCTATT---	GTGTTGTATGA	AACCCCAAAA	GCCAATGTGC	GTGGAGTCCT	TACCCGATTA	TCCACCTCTT	GGTCGGTTTG	CTGTGAGGGA	CATGCGGCCAA	[1134]
109_EF1a_Euglena_viridis	CTGGTGATGC	TGCCATC---	GTCTCATGA	AGCCCCAGAA	GCCTATGTGC	GTGGAGTCCT	TACAGACTA	---	---	---	---	[1172]
110_EF1a_Eutreptia_pertyi	CTGGTGATGC	TGCCATC---	GTGCTGTATGA	AGCCTCAGAA	GCCTATGTGC	GTGGAGTCCT	TACCCGACTA	CCCCCCCCCT	GGCCC-----	---	---	[1229]
111_EF1a_Eutreptiella_pomquetensis	CTGGTGATGC	TGCCATC---	GTCTGTATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TACAGGACTA	CCCCCCCTT	GGCCC-----	---	---	[839]
112_EF1a_Lepocinclis_acus_var._major	CTGGTGATGC	TGCCATT---	GTGCTGTATGA	AGCCTCAGAA	GCCAATGTGC	GTGGAGTCGT	TACCCGACTA	CCCCCCCT-	---	---	---	[1223]
113_EF1a_Lepocinclis_ovum	CTGGCGATGC	TGCCATC---	GTGCTGTATGA	AGCCCCAGAA	GCCAATGTGC	GTGGAGTCCT	TCAGTGATTA	CCCTCCGTTG	GGACGGTTTG	CYGTCCGTGA	TATGCGGCAR	[1134]
114_EF1a_Lepocinclis_spyrogyroides_clone_38	CCGGTGATGC	CGCCATT---	GTGCTGTATGA	AGCCCCAGAA	GCCAATGTGC	GTGGAGTCCT	TCACCGACTA	CCCCCCCTTG	GACC-----	---	---	[541]
115_EF1a_L. spyrogyroides_clones_1_&_4	CTGGTGATGC	GGCCATT---	GTGCTCATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCGT	TACCCGACTA	CCCCCCCTTA	GGC-----	---	---	[1227]
116_EF1a_Lepocinclis_steinii	CTGGTGATGC	TGCCATC---	GTGCTGTATGA	AGCCCCAGAA	GCCAATGTGC	GTGGAGTCCT	TCAGTGATTA	CCCACTCTG	GGACGGTTTG	CCGTGCGTGA	CATGCGGCAG	[1134]
117_EF1a_Lepocinclis_tripteris	CTGGTGATGC	AGCCATC---	GTGCTGTATGA	AGCCTCAGAA	GCCCATGTGC	GTGGAGTCCT	TYACCGACTA	CCCCCCCTT	GGACC-----	---	---	[1229]
118_EF1a_Monomorphina_aenigmatica	CTGGTGATGC	TGCTATT---	GTCTGTATGA	AGCCTCAGAA	GCCCATGTGC	GTGGAGTCCT	TACCCGACTA	CGCCCCCTTA	GGCCC-----	---	---	[1229]
119_EF1a_Monomorphina_pyrum	CTGGTGATGC	GGCCATT---	GTGCTCATGA	AGCCTCAGAA	GCCCATGTGC	GTGGAGTCGT	TACAGGACTA	CCCCCCCTC	GGACC-----	---	---	[1229]
120_EF1a_Monomorphina_sp.	CTGGTGATGC	CGCCATT---	GTGCTCATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCAT	TCAMCGAYTA	SCCCCCCTT	GGACC-----	---	---	[1228]
121_EF1a_Phacus_orbicularis	CTGGAGATGC	TGCCATC---	GTCTGTATGA	AGCCCCAGAA	GCCCA-----	---	---	---	---	---	---	[1147]
122_EF1a_Phacus_similis	CTGGGGACGC	CGCTATT---	GTGCTGTATGA	AGCCGCGAGAA	GCCCATGTGC	GTGGAGTCGT	TNACCGACTA	CCCCCCCCCT	GGGCGCTTTG	CTGTGCGTGA	CATGCGGCAG	[1223]
123_EF1a_Strombomonas_acuminata	CTGGGGACGC	TGCCATT---	GTGCTGTATGA	AGCCGCGAGAA	GCCCATGTGC	GTGGAGTCCT	TACAGGACTA	CCCCCCACTG	GGGCGCTTTG	CTGTGCGTGA	CATGCGGCAG	[1257]
124_EF1a_Strombomonas_borsteniensis	CTGGTGATGC	TGCCATC---	GTCTGTATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TCAGTGATTA	TCCCCCTCTG	GGACGGTTTG	CTGTCCGTGA	CATGCGGCCAA	[1134]
125_EF1a_Strombomonas_napiformis	CTGGTGACGC	AGCCATC---	GTGCTCATGA	AGCCACAGAA	GCCCATGTGC	GTGGAGTCCT	TACAGGACTA	CCCCCCCTT	GGGCC-----	---	---	[1229]
126_EF1a_Strombomonas_ovalis	CTGGTGATGC	TGCTATC---	GTCTGTATGA	AGCCACAGAA	GCCCATGTGC	GTGGAGTCCT	TACCCGACTA	TCCTCCTCTG	GGACGGTTTG	CTGTCCGTGA	CATGCGACAA	[1254]
127_EF1a_Strombomonas_triquetra	---	---	---	---	---	---	---	---	---	---	---	[891]
128_EF1a_Trachelomonas_conradi	CCGGTGATGC	TGCCATT---	GTGCTGTATGA	AGCCCCAGAA	GCCCATG---	---	---	---	---	---	---	[1191]
129_EF1a_Trachelomonas_echinata	CTGGTGATGC	TGCCATT---	GTGCTGTATGA	AGCCACAAAA	GCCAATGTGC	GTGGAGTCAT	TCACCGACTA	CCCCCCNNNT	TC-----	---	---	[1226]
130_EF1a_Trachelomonas_lefevrei	CAGGTGATGC	CGCCATT---	GTGCTCATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCAT	TACCCGACTA	CCCCCCCTT-	---	---	---	[1223]
131_EF1a_Trachelomonas_plantonica	CAGGTGATGC	CGCCATT---	GTGCTCATGA	AGCCCCAGAA	GCCCATGTGC	GTGGAGTCCT	TCAGTGATTA	CCCACTCTT	GGACGGTTTG	CTGTGCGTGA	CATGAGGCAG	[1134]
132_EF1a_Trachelomonas_similis	CTGGAGATGC	CGCTATT---	GTCTGTATGA	AGCCACAAAA	GCCCATGTGC	GTGGAGTCCT	TCAGTGATTA	CCCACTCTT	GGACGGTTTG	CTGYCCGTGA	CATGCGGCAR	[1134]

	1550	1560	1570	1580	1590	1600	1610	1620	1630	1640	1650	
133_EFL1a_Trachelomonas_volvocina	CCGGAGATGC	GCCATT---	GTGCTCATGA	AGCCCCAGAA	GCCCATGTGT	GTGAGTTCGT	TCACCGATTG	CCCGCTCT-	-----	-----	-----	[1223]
134_EFLenvironmental_sample_EU810232	GCAACGAGAT	GCCCCAG---	CTCGTCTTTG	AGCCCCAGCA	GCCGTTCTGC	GTGACACGCT	TCAAGAGCTG	CGAGGGACTC	GGCCGTGTGC	CCATCATGGA	GGGCAACGGA	[1347]
135_EFLCreolimax_fragrantissima												[994]
136_EFLSphaeroforma_arctica	CTGCCGATGT	CAAA-										[1255]
137_EFLMonosiga_brevicollis	CCAACGAGGT	TGCCGAG---	GTGCTCATGG	CCCCCAGCA	GCCCTTCTGC	GTGCACACCT	TCAAGACCTG	CGAGGGTCTT	GGCCGTGTGC	CCATCATGGA	GGGTAACCTG	[1347]
138_EFLXenopus_laevis	CCAACGAGAT	GCCCCGAG---	TGCACCTTCG	AGCCCCATGCA	GCCCTTCTGC	GTGCACCTGT	TCAAGAACTG	CGAGGGTCTG	GGCCGCTGCG	CGCTCATGGA	GGGCAACTCC	[1347]
139_EFLSpizellomyces_punctatus	CCAACGAGGC	TGCTGAG---	GTGCTGTTTCG	AGCCCCATCA	ACCCCTTGTG	GTGCACCTCT	TCAAGAACTG	CGAGGGTCTC	TCCCGATTG	CCTTCCTTGA	CGGTAACCTCT	[1353]
140_EFLGaertneriomyces_semiglobiferous												[506]
141_EFLAllomyces_macrogynus	CCAACGAGGC	GCCCCGAG---	GTTGTCTTCG	AGCCCCATCA	GCCGCTCGTT	GTGCACCTGT	TCCAGAACTG	CGAGGGTCTG	TCGCGATTG	CCTTCCTCGA	TGGCAACACG	[1353]
142_EFLBlastocladiella_emersonii	CCAACGAGGC	CGCTGAG---	GTGCTCTTCG	AGCCCCATCA	GCCCTCTGTC	GTGCACACCT	TCGGCAACTG	CGAGGGTCTG	TCGCGCATTG	CCTTCCTCGA	CGGCAACACC	[1353]
143_EFLCatenophylctis_sp.												[506]
144_EFLCoelomomyces_stegomyiae												[506]
145_EFLEntomophthora_muscae												[503]
146_EFLBasidiobolus_ranarum												[503]
147_EFLConidiobolus_coronatus												[503]
148_EFLTetraselmis_tetrathele	CCAACGAGGT	CTCGACC---	ATCGTGTTCG	AGCCCTGCCA	GCCCTGTGGT	GTGGACTCGT	ACAAGAACTG	CGAGGGCTG	TCCCGCATTG	CGTTCTTGGG	CGGCAACACC	[1296]
149_EFLMesostigma_viride												[879]
150_EFLOstreococcus_tauri	CCAACGAAGT	CGCCGAG---	TGCGTCTTCG	AACC6GTGCA	CCC6CTCGTC	GTGACACTCT	TCAAGAACTG	CGAGGGTCTT	TCCCGCATCG	CCTTCCTTGA	CGGTAACACC	[1356]
151_EFLNephroselmis_olivacea												[957]
152_EFLMicromonas_pusilla	CCAACGAGGT	TGCCGAG---	GCCGCTCTTCG	AGCCCCAGCA	CCCCCTTCTG	GTGCACCACT	TCAAGGCTTG	CGACGGTCTC	GCCCGCATCG	CGTTCTTGA	CGGCAACACC	[1305]
153_EFLParachlorella_kessleri												[942]
154_EFLHelicosporidium_sp.	CCAACGAGAT	GGCTGAG---	GTGCTGTTTCG	AGCCCTGCCA	GCCCTGTGGT	GTGCACCTCT	TCAAGAACTG	CGAGGGTCTC	TCCCGCATTG	CCTTCCTTGA	CGGCAACACC	[1368]
155_EFLGonium_pectorale	CCAACGAGAT	GGCCGAG---	GTGGTCTTCG	AGCCCCACCA	GCCCTTGATT	GTGGAGTCGG	CAAAGAGCTG	CGAGGGCTG	TCGCGCATTG	CTTTCTTGGG	CGGCAACACC	[1350]
156_EFLChlamydomonas_reinhardtii	CCAACGAGAT	GGCTGAG---	GTGCTCTACG	AGCCCCACCA	GCCCTGATC	GTGGACTCCT	TCAAGAACTG	CGAGGGCTG	TCCCGCATTG	CCTTCCTTGA	CGGCAACACC	[1350]
157_EFLScenedesmus_obliquus	CCAACGAGAT	GGCGCAG---	GTGACCTTCG	AGCCCGTGCA	GCCCTGTGTC	GTGGACAGCT	TCAAGTCTCT	CGAGGGCTG	TCCCGCATTG	CCTTCCTTGA	CGGCAACACC	[1350]
158_EFL"Oryza_sativa"	CCAACGAGAT	GGCTGAG---	GTGGTATTTC	AGCCATGCCA	GCCCTGTGGT	GTGGAGCTCT	TCAAGAGCTG	TGAGGGCTG	TCCCGCATTG	CCTTCCTTGA	CGGCAACACC	[1350]
159_EFLChlorococcum_sp.	CCAACGAGAT	GGCTGAG---	ATTGTGTTCA	CTCCCTGCCA	GCCCTTGTG	GTGGACTCCT	TCAAGAACTG	CGAGGGTCTC	TCCCGATTG	CGTTCTTGA	CGGTAATACC	[1299]
160_EFLOchlochaete_hystrix	CCAACGAGAT	GGCTGAG---	GTGGTGTTCG	AGCCCTGCCA	GCCCTGTGTC	GTGGACTCCT	TAAAGAACTG	C-				[1269]
161_EFLUlva_intestinalis	CCAACGAGAT	GGCACAG---	ATCGTCTTCA	AGCCCTGCCA	GCCCTTGTG	GTGCACACTT	TCAAGAACTG	CGAGGGTCTG	TCCCGCATTG	CCTTCCTTGA	CGGCAACACC	[1299]
162_EFLUrospora_sp.	CCAACGAGAT	GGCGCAG---	GTGCTCTTCA	AGCCGACGCA	GCCG-							[1086]
163_EFLGymnochlorella_stellata_ppc	ATGGAGAGAT	GGCTGAG---	GTGCTGTTTC	AGCCGATGTT	TGGAATTCGTG	CGGAGGCTG	TAAAGAGATG	CGAAGGCTT	GCACGATTG	CGTTCTTAGA	CGGAAACGGT	[1386]
164_EFLLotharella_amoebiformis_ppc	ACGGCGAAAT	GGCTGAA---	GTGGTGTTCG	AGCCGATGTT	CGGGTTCGTT	ACCGAGACTG	TCAAGACATTG	TGAAGGCTC	GCCCGGATCG	CCTTCCTCGA	CGGAAACGGA	[1131]
165_EFLChlorarachnion_reptans_ppc	ATGGGGATAT	GGCTGAA---	GTGGTGTTCG	AACCATGCTT	CGGATTCGTA	GCTGAGGGAT	TCAAAAGTTG	TGAAGGCTT	GCCCGATTG	CCTTCCTTGA	TGGCAACGGT	[1158]
166_EFLChondrus_crispus	CTAAGCAAGT	GGCTGAA---	ATTGTTTTTC	CTCCGACGCA	ACCAATTGTT	GCTGAGAGAT	TCAAGAACTG	CGAGGGCTT	GGCGGTGTT	CCATTATGGA		[1337]
167_EFLGracilaria_changii	CTAATGAAGT	TGCTGAG---	ATGCTCTTTG	AGCCGACGCA	GCCATTGCTT	GTGAGGGAT	TCAAGGCTG	CGAAGGACTC	GGTCTGTAG	CTATTATGGA	GGGAAACACG	[1347]
168_EFLCryptomonas_ovata	CGAACGAGGT	TGCCAG---	GCACAATTCC	TGCCGATGAC	TCCCTTGTG	TGCGACAACT	TCAAGAACTG	CGAGGGCTC	TCACGGATTG	CATTTCAGA	CGGAAAC---	[1305]
169_EFLGuillardia_theta	CCAACGAGGT	TGCCAG---	GCTCAGTTCC	AACCCATGAC	TCCCTTGTG	TGCGACACTT	TCAAGAACTG	CGAGGGCTG	AGCCGATTG	CCTTCCTCGA	CGGAAACACT	[1353]
170_EFLRhodomonas_salina	CCAACGAGGT	CGCCAG---	GCCAGTTCCG	AGCCGATGAC	TCCCTTGTG	TGCGACACTT	TCAAGAACTG	CGAGGGTCTC	TCCCGCATTG	CTTTCTTGA	CGGAAACACC	[1350]
171_EFLPavlova_lutheri												[1206]
172_EFLChrysochromulina_sp.												[1017]
173_EFLIsochrysis_galbana	CCAACGAGAT	GGCGCAA---	GTCTGTTTC	GCCCGACGCA	GCCGCTCGTC	GTGATTCCT	TCAAGAACTG	CGAGGGCTC	TCGCGCATTG	CCTTCCTCGA	CGGCAACACC	[1350]
174_EFLEmiliana_huxleyi	CGAACGAGAT	GGCGCAG---	GTCTCTTTC	GCCCGACGCA	GCCGCTCGTC	GTGCAGCTGT	TCAAAAACTG	CGAGGGCTG	TCCCGCATCG	CCTTCCTCGA	CGGCAACACC	[1356]
175_EFLAchnanthes_kuwaitensis	AGTACGAATC	TGCTGAG---	GTGCAATTTC	AGCCCACTGC	CCCGCTCTTT	GTGAGGCTAT	TCGAGACTTG	CGCCGCTCTT	GGCCGATTG	CCGTCATGGA	TTCCAACCTG	[1257]
176_EFLAsterionella_glacialis	AATTTCGAGAA	GGCCGAG---	GTGCAAGTTC	AACCCACTCA	GCTCTTTTTC	CTTGAACCTT	TGAGAACTG	TGCTGCTCTT	GGACGATTG	CTGTTATGGA	TTCCAACCTG	[1257]
177_EFLDetonula_confervacea	AATACGAGAA	TGCTGAG---	GTTGAATTTC	TGCTTACTGC	TCTCTCTTTC	GTGAGGCTT	TCGAGGCTTG	CGCCGCTCTT	GGACGATTG	CTGTCATGGA	CTCCAACAGA	[1257]
178_EFLThalassiosira_pseudonana	AGTTTCGAGAA	GGCCGAG---	GTTGAGTTTC	AGCCCAAGTCA	GCTCTTTTTC	GTGAGGCTT	TCGAGAACTG	TGACGCTCTT	GGACGATTG	CCGTAATGGA	CTCCAACAGA	[1320]
179_EFLPerkinsus_marinus	CTAACGAAAT	GGCTGAA---	GTGCTATTTC	AACCTATTCA	GCTCTTGTG	GTGCACCTCT	TCAAGAACTG	TGAGGGTTG	TCGCGTGTG	CTCTGATGGA	GGGTAACGGG	[1365]
180_EFLOxyrrhis_marina	CCAACGAGAT	GGCTGAG---	GTGCTGTTCA	GCCCGATGCA	GCCGTTGGTG	GTGGACACTT	TCAAGAACTG	CGAGGGCTG	TCCCGCATTG	CTTTCTTGA	TGGCAATGGT	[1359]
181_EFLKarlodinium_micrum	CAAACGAAAT	GGCGGAG---	GCCCTCTTTC	AACCCACAGCA	GCCCTCGTG	TGCGACACTT	TCAAGAACTG	CGAGGGCTC	TCGCGTGTG	CTTTCTTGA	CGGAAATGGC	[1353]
182_EFLAmphidinium_carterae	CCAACGAGAT	GGCCGAG---	TGCACCTTTC	AGCCACAGCA	GCCACTGGTG	TGCGACAGCT	TCAAGAACTG	CGAAGGCTG	TCTCGGCTG	CCTTCATGGA	TGGCAACG--	[1351]
183_EFLHeterocapsa_triquetra	CGAACGAGAT	GGCTGAG---	GCCGCTTTC	AGCCGACAGCA	GCCGCTTGTG	TGCGACACTT	TCAAGAACTG	CGAGGGCTT	TCCCGTGTG	CTTCATGGA	GGGTAACGGT	[1353]
184_EFLPlanoglabratella_operculis												[961]
185_EFLReticulomyxa_filosa	CCGGTGACCA	AGCTGAA---	GTGCTATTTC	CTCCAATAAT	GCCATTAGTC	GTGACATCCT	TTGATGAGTG	CAAAACATTG	GGACGAATG	CTGCCATGGA	CTGAACTCG	[1293]
186_EFLLotharella_amoebiformis	CCGGTGACCA	GTGCGAG---	ATTGAGTTTC	AGCCCGAGGG	ACCTCTCTCT	GTCTCTCCT	TCAAGGAGTG	CGAGGGCTT	GGACGATTG	CTGCTTGA	GTGCAACTCG	[1458]
187_EFLGymnochlorella_stellata	AGGGAGACGT	AGCTGAA---	ATCGAGTTTC	AGCCCTAGAT	GCCTATGCT	GTCTTCACT	TAAAGCAGTG	TGAGGGTCTT	GGACGTGTT	CTGTTCTCGA	GTCTAACTCT	[1431]
188_EFLChlorarachnion_reptans	AGGGGACCTC	CGCTGAA---	CTTGAGTTTC	AACCCAAAT	GCCATTGCTT	ATCTCTACT	TACGCACTG	TGAGGACTT	GGTCTGTTG	CGCTTCTGA	GTCTAACTCT	[1362]
189_EFLDiplonema_ambulator	CCAACGAGAT	GGCCGAG---	GTGCTGTTTC	CCCCCAGCA	GCCCTCTGTC	GTGCAGACT	TCAAGAACTG	CGAGGGTCTC	TCCCGTGTG	CCATCATGGA	GGGGAACGGT	[951]
190_EFLDiplonema_papillatum	CCAACGAGAT	GGCTGAG---	GTTGTGTTTC	GCCCGACAGCA	GCCGCTTGTG	GTGCAGACT	TCAAGAACTG	CGAGGGTCTC	TCGCGTGTG	CCATCATGGA	GGGTAACGGT	[1323]
191_EFLRhynchopus_euleeides	CCAACGAGAT	GGCTGAG---	GTGCTGTTTC	CCCCCGTGA	GCCCTGTGTT	GTGCACACT	TCAAGAACTG	CGAGGGTCTG	TCCCGTGTG	CCATCATGGA	GGGCAACGGT	[966]
192_EFLPetalomonas_cantuscygni												[924]
193_EFLTrypanoplasma_borreli												[1017]
194_EFLNeobodo_salienis	AGGGGACAT	GTGCGAG---	CTGCTTGGG	AGCCGACGCG	CCCGTTCGTC	GTGCAGACT	TCAAGAGC--					[1278]

[illegible]

		1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760
133	EFL1a Trachelomonas volvocina	---	---	---	---	---	---	---	---	---	---	---
134	EFL environmental sample EU810232	GTGGTCATGC	TCGGAAGGT	CGTCAGCGTC	ACCTTCAAGG	AG---	---	---	---	---	---	---
135	EFL Creolimax fragrantissima	---	---	---	---	---	---	---	---	---	---	---
136	EFL Sphaeroforma arctica	---	---	---	---	---	---	---	---	---	---	---
137	EFL Monosiga brevicollis	GTGTGCATGC	TGGGCAAGGT	CACCAAGGTC	GAGCGTGCT-	---	---	---	---	---	---	---
138	EFL Xenopus laevis	GTGGTCATGC	TGGGCAAGGT	CGTGACTGTC	GAGTTCCTG	---	---	---	---	---	---	---
139	EFL Spizellomyces punctatus	GCTGTCACTG	TGGGCAAGGT	AACGCAAGTC	GAATACAAGG	AT---	---	---	---	---	---	---
140	EFL Gaertneriomyces semiglobiferous	---	---	---	---	---	---	---	---	---	---	---
141	EFL Allomyces macrogynus	GCCGTCATGC	TCGGCAAGGT	TGTCGCTGTG	GAGCACAAG-	---	---	---	---	---	---	---
142	EFL Blastocladia emersonii	GCCGTCATGC	TGGGCAAGGT	CGTCGCTGTT	GAGCACAAG-	---	---	---	---	---	---	---
143	EFL Catenophylctis sp.	---	---	---	---	---	---	---	---	---	---	---
144	EFL Coelomomyces stegomyiae	---	---	---	---	---	---	---	---	---	---	---
145	EFL Entomophthora muscae	---	---	---	---	---	---	---	---	---	---	---
146	EFL Basidiobolus ranarum	---	---	---	---	---	---	---	---	---	---	---
147	EFL Conidiobolus coronatus	---	---	---	---	---	---	---	---	---	---	---
148	EFL Tetraselmis tetrathele	GCCGTAATGA	TGGGCAAGGT	CATGCTGCC	ACCATGAGCG	CC---	---	---	---	---	---	---
149	EFL Mesostigma viride	---	---	---	---	---	---	---	---	---	---	---
150	EFL Ostreococcus tauri	GCGGTCATGC	TCGGTAAGAT	CACGAACGTC	GAGTACAAGG	CCATC---	---	---	---	---	---	---
151	EFL Nephroselmis olivacea	---	---	---	---	---	---	---	---	---	---	---
152	EFL Micromonas pusilla	GCCGTCATGC	TTGGCAAGAT	CACCAAGGTC	CAGTTACAAGG	AGGACAAG--	---	---	---	---	---	---
153	EFL Parachlorella kessleri	---	---	---	---	---	---	---	---	---	---	---
154	EFL Helicosporidium sp.	GCCGTCATGC	TTGGCAAGGT	GGTGTCCACC	ACCTCCAAG-	---	---	---	---	---	---	---
155	EFL Gonium pectorale	GCGGTCATGC	TCGGCAAGGT	CACGACGCTG	CTCTTCAAG-	---	---	---	---	---	---	---
156	EFL Chlamydomonas reinhardtii	GCCGTCATGC	TCGGCAAGGT	GGTGTGCGTC	ACCCACAAG-	---	---	---	---	---	---	---
157	EFL Scenedesmus obliquus	GCTGTCATGC	TCGGCAAGGT	GGTGACCGTC	ACCTTCAAG-	---	---	---	---	---	---	---
158	EFL "Oryza sativa"	GCTGTCATGC	TTGGCAAGGT	CGTGCGTGTG	ACACACAAG-	---	---	---	---	---	---	---
159	EFL Chlorococcum sp.	GCTGTCATGC	TTGGCAAGGT	CGTGAAGGTC	ACCCACAAG-	---	---	---	---	---	---	---
160	EFL Ochlochaete hystrix	---	---	---	---	---	---	---	---	---	---	---
161	EFL Ulva intestinalis	---	---	---	---	---	---	---	---	---	---	---
162	EFL Urospora sp.	GCTGTCATGC	TTGGCAAGGT	GGTGAAGGTC	ACCCACAAG-	---	---	---	---	---	---	---
163	EFL Gymnochlorella stellata_ppc	GTGCTTATGC	TTGGCAAGGT	AACAAGAGCT	GTTCCAAAAA	GCTTTGGATC	AGATAAG---	AAAGATAAAG	GTGGCAAAGA	TGGA-----	GATAAGAAGG	GCAAGGACGA
164	EFL Lotharella amoebiformis_ppc	GTGCTCATGC	TTGGAAAGGT	TATAAGAGCT	ATCTCGAAGC	AGTTTTGCTGA	TCCCAAG---	AAGAAGGCCG	GTCTCTGGTGA	CAAGAAGAAA	GACGACAAGA	AGAAAGACGA
165	EFL Chlorarachnion reptans_ppc	GTGTGTTATGC	TTGGAAAGGT	TACAAAGACG	CAACCAAAAGT	CATATGGGTAA	CAAAGGAGAC	AAGAAAGGAG	ACAAGAAAGG	CGACGGGAAG	GATGATAAAA	AAGACGACAA
166	EFL Chondrus crispus	---	---	---	---	---	---	---	---	---	---	---
167	EFL Gracilaria changii	GTGTGTTATGC	TCGGAAAAAGC	CATCAAGGTC	GAGTTCTACA	CGG---	---	---	---	---	---	---
168	EFL Cryptomonas ovata	---	---	---	---	---	---	---	---	---	---	---
169	EFL Guillardia theta	GTATTGATGC	TTGGCAAGGT	TATTGCAACA	GTTGCTCGCG	ATGATGGTGG	TGCTAAGAAG	AAA-----	---	---	---	---
170	EFL Rhodomonas salina	GTGATGATGC	TTGGCAAGGT	CATTGCCACC	GTTGCCCGTG	AGGACGCTGG	CAAGAAGAAG	---	---	---	---	---
171	EFL Pavlova lutheri	---	---	---	---	---	---	---	---	---	---	---
172	EFL Chrysochromulina sp.	---	---	---	---	---	---	---	---	---	---	---
173	EFL Isochrysis galbana	GCCGTCATGC	TTGGCAAGGT	CGTGTCCGTC	ATTGCCAAGC	AGGACATGCC	GCCGGAGGAG	GACAAGAAGA	AGAAGAAG--	---	---	---
174	EFL Emiliana huxleyi	GCCGTCATGC	TTGGCAAGGT	CGTGTCCGTC	ATCCAGAAGA	AGGACATGCC	GCCGGAGGAG	GGCAAGAAGA	AGAAG-----	---	---	---
175	EFL Achnanthes kuwaitensis	TTGAAGATGC	TTGGAAGAGT	CACC6GGGTT	GTCTCCAGCA	CCGCA----	---	---	---	---	---	---
176	EFL Asterionella glacialis	CTTAAGATGC	TTGGAAGAGT	CACACACACT	GAATCA-----	---	---	---	---	---	---	---
177	EFL Detonula confervacea	TTGAAGATGT	TGGGTAAAGT	CACC6GCACC	ACCACCGTTG	AA-----	---	---	---	---	---	---
178	EFL Thalassiosira pseudonana	TTGAAGATGC	TTGGAAGAGT	CACGTGGAGTT	ACTTCTGTTT-	---	---	---	---	---	---	---
179	EFL Perkinsus marinus	GTGGTCATGC	TTGGCAAGAT	TGTCCTCCACC	GTGGCGAAGG	ACGTGAAG--	---	---	---	---	---	---
180	EFL Oxyrrhis marina	GCGGTCATGC	TGGGCAAGAT	CGTGTCCAAS	GTCGACAAGG	GTGTCTGATGA	TGGCAAGGCC	GGGAAGGCTG	GTAAGAAGAA	G-----	---	---
181	EFL Karlodinium micrum	GTAGTCATGC	TTGGGAAGGT	CATCTCTTGC	GAGAAGAAAG	ACGAGAACGA	CAAAGCGGCG	GGCAAGAAGA	AG-----	---	---	---
182	EFL Amphidinium carterae	---	---	---	---	---	---	---	---	---	---	---
183	EFL Heterocapsa triquetra	GTGTGATGC	TTGGCAAGGT	CATCTCGTGT	GAGCGTAAGG	AGGAGGGCGA	CAAGGGTGCG	AAGAAGAAG-	---	---	---	---
184	EFL Planoglabratella operculis	---	---	---	---	---	---	---	---	---	---	---
185	EFL Reticulomyxa filosa	TTGATTATGT	TGGGAAAAGT	TGTGTCAGTT	GAATACAAG-	---	---	---	---	---	---	---
186	EFL Lotharella amoebiformis	CTCGTTATGC	TTGGAAGATG	CAACTCGGTT	GTCACACCCG	AAATCAAG--	---	---	---	---	---	---
187	EFL Gymnochlorella stellata	CTCGTTATGA	TGGTAAAGAT	CCAAAGTGCA	GAGCACGGCG	CTGTCAAGGA	TC TTGCA---	---	---	---	---	---
188	EFL Chlorarachnion reptans	CTTGTTATGA	TTGGCAAGAT	CATTGCAGCA	GAGCAGACAG	CGGTTGCCGA	CCGCACA----	---	---	---	---	---
189	EFL Diplonema ambulator	GTCGTCATGC	TTGGCAAGGT	CACCAAGACC	TCCGAGAAGG	TCGTGCGCCG	CAAGAAGAAG	---	---	---	---	---
190	EFL Diplonema papillatum	GTGTGTCATGC	TTGGCAAGGT	CGTCAAGACC	GCCGAGAAGG	CCAAGGACGA	CGGCAAGGGC	AAGAAGAAG-	---	---	---	---
191	EFL Rhynchopus euleoides	GTGGTCATGC	TTGGCAAGGT	CACCAAGACC	ACCGAGAARG	TTGCCGGGGA	CAAGAAGGGC	AAGAAG-----	---	---	---	---
192	EFL Petalomonas cantuscygni	---	---	---	---	---	---	---	---	---	---	---
193	EFL Trypanoplasma borreli	---	---	---	---	---	---	---	---	---	---	---
194	EFL Neobodo saliens	---	---	---	---	---	---	---	---	---	---	---

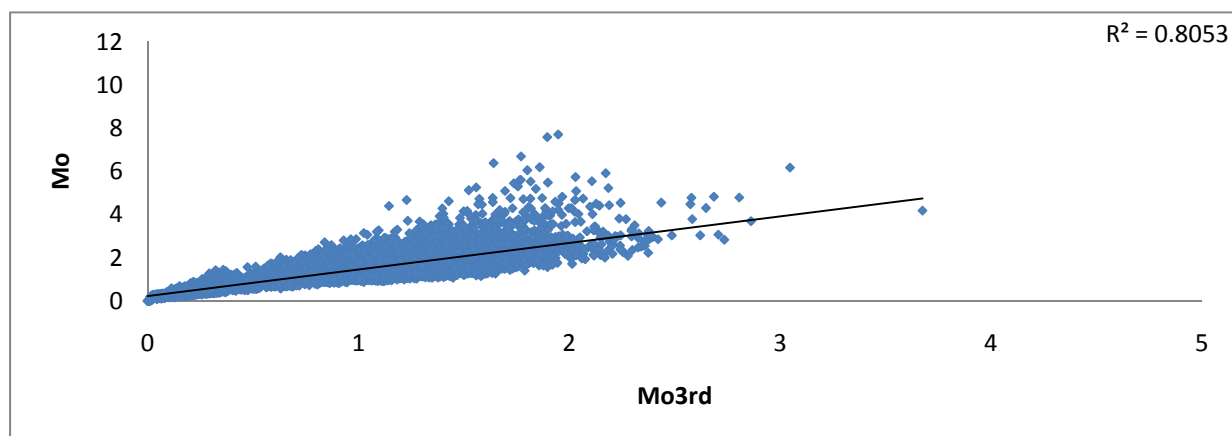
	1770	1780	1790	1800	
001_aEF1a_Desulfurococcus_mobilis	-----	-----	-----	---	[1314]
002_aEF1a_Sulfolobus_solfataricus	-----	-----	-----	---	[1305]
003_aEF1a_Thermoplasma_acidophilum	-----	-----	-----	---	[1272]
004_aEF1a_Halobacterium_sp.	-----	-----	-----	---	[1263]
005_aEF1a_Methanocaldococcus_jannaschii	-----	-----	-----	---	[1284]
006_EF1a_Acanthamoeba_culbertsoni	-----	-----	-----	---	[1245]
007_EF1a_Neoparamoeba_branchipila	-----	-----	-----	---	[1201]
008_EF1a_Planoprotostelium_aurantium	-----	-----	-----	---	[1203]
009_EF1a_Physarum_polycephalum	-----	-----	-----	---	[1203]
010_EF1a_Amoebidium_parasiticum	-----	-----	-----	---	[867]
011_EF1a_Capsaspora_owczaraki	-----	-----	-----	---	[1393]
012_EF1a_Monosiga_ovata	-----	-----	-----	---	[1284]
013_EF1a_Salpingoeca_amphoridium	-----	-----	-----	---	[1290]
014_EF1a_Ceanorhabditis_elegans	-----	-----	-----	---	[1389]
015_EF1a_Drosophila_melanogaster	-----	-----	-----	---	[1389]
016_EF1a_Ornithorhynchus_anatinus	-----	-----	-----	---	[1389]
017_EF1a_Xenopus_laevis	-----	-----	-----	---	[1383]
018_EF1a_Saccharomyces_cerevisiae	-----	-----	-----	---	[1374]
019_EF1a_Neurospora_crassa	-----	-----	-----	---	[1380]
020_EF1a_Cryptococcus_neoformans	-----	-----	-----	---	[1377]
021_EF1a_Schizophyllum commune	-----	-----	-----	---	[1380]
022_EF1a_Monoblepharis_insignis	-----	-----	-----	---	[1169]
023_EF1a_Rhizophlyctis_harderi	-----	-----	-----	---	[1169]
024_EF1a_Harpochytrium_sp.	-----	-----	-----	---	[1077]
025_EF1a_Absidia_glauca	-----	-----	-----	---	[1374]
026_EF1a_Smittium_simulii	-----	-----	-----	---	[1281]
027_EF1a_Chara_australis	-----	-----	-----	---	[1293]
028_EF1a_Physcomitrella_patens	-----	-----	-----	---	[1341]
029_EF1a_Acetabularia_acetabulum	-----	-----	-----	---	[1290]
030_EF1a_Chaetomorpha_coliformis	-----	-----	-----	---	[1137]
031_EF1a_Arabidopsis_thaliana	-----	-----	-----	---	[1347]
032_EF1a_Oryza_sativa	-----	-----	-----	---	[1341]
033_EF1a_Cyanidioschyzon_merolae	-----	-----	-----	---	[1350]
034_EF1a_Porphyrta_yezoensis	-----	-----	-----	---	[1347]
035_EF1a Plasmodium_falciparum	-----	-----	-----	---	[1329]
036_EF1a_Cryptosporidium_parvum	-----	-----	-----	---	[1305]
037_EF1a_Tetrahymena_pyriformis	-----	-----	-----	---	[1305]
038_EF1a_Stentor_coeruleus	-----	-----	-----	---	[1227]
039_EF1a_Phytophthora_parasitica	-----	-----	-----	---	[1329]
040_EF1a_Mallomonas_rasilis	-----	-----	-----	---	[1242]
041_EF1a_Phaeodactylum_tricornutum	-----	-----	-----	---	[1317]
042_EF1a_Thalassiosira_pseudonana	-----	-----	-----	---	[1290]
043_EF1a_Malawimonas_jakobiformis	-----	-----	-----	---	[1185]
044_EF1a_Malawimonas_californiana	-----	-----	-----	---	[1338]
045_EF1a_Trimastix_pyriformis	-----	-----	-----	---	[1212]
046_EF1a_Dinenympha_exilis	-----	-----	-----	---	[1185]
047_EF1a_Monocercomonoides_sp.	-----	-----	-----	---	[1344]
048_EF1a_Streblo mastix_strix	-----	-----	-----	---	[1194]
049_EF1a_Histomonas_meleagridis	-----	-----	-----	---	[1320]
050_EF1a_Trichomonas_vaginalis	-----	-----	-----	---	[1311]
051_EF1a_Hexamita_inflata	-----	-----	-----	---	[1254]
052_EF1a_Spironucleus_vortens	-----	-----	-----	---	[1200]
053_EF1a_Andalucia_incarcerata	-----	-----	-----	---	[1203]
054_EF1a_Seculamonas_ecuadoriensis	-----	-----	-----	---	[1341]
055_EF1a_Reclinomonas_americana	-----	-----	-----	---	[1356]
056_EF1a_Histiona_aroides	-----	-----	-----	---	[1335]
057_EF1a_Jakoba_libera	-----	-----	-----	---	[1329]
058_EF1a_Stachyamoeba_lipophora	-----	-----	-----	---	[1299]
059_EF1a_Naegleria_gruberi	-----	-----	-----	---	[1206]
060_EF1a_Naegleria_andersoni	-----	-----	-----	---	[1215]
061_EF1a_Acrasis_rosea	-----	-----	-----	---	[1203]
062_EF1a_Sawyeria_marylandensis	-----	-----	-----	---	[1305]
063_EF1a_Bodo_saltans	-----	-----	-----	---	[1108]
064_EF1a_Rhynchobodo_sp.	-----	-----	-----	---	[1102]
065_EF1a_Rhynchomonas_nasuta	-----	-----	-----	---	[1214]
066_EF1a_Dimastigella_trypaniformis	-----	-----	-----	---	[1126]

	1770	1780	1790	1800	
067_EF1a_Perkinsiella_amoebae	-----	-----	-----	-----	[915]
068_EF1a_Leishmania_major	-----	-----	-----	-----	[1347]
069_EF1a_Leishmania_infantum	-----	-----	-----	-----	[1347]
070_EF1a_Leishmania_braziliensis	-----	-----	-----	-----	[1347]
071_EF1a_Leishmania_tarentolae	-----	-----	-----	-----	[1032]
072_EF1a "L. tarentolae" FJ807257	-----	-----	-----	-----	[915]
073_EF1a_Leptomonas_costaricensis	-----	-----	-----	-----	[594]
074_EF1a_Leptomonas_bifurcata	-----	-----	-----	-----	[832]
075_EF1a_Leptomonas_podlipaevi	-----	-----	-----	-----	[814]
076_EF1a_Herpetomonas_muscarum	-----	-----	-----	-----	[1144]
077_EF1a_Herpetomonas_pessoai	-----	-----	-----	-----	[771]
078_EF1a_Blastocrithidia_culicis	-----	-----	-----	-----	[766]
079_EF1a_Trypanosoma_brucei	-----	-----	-----	-----	[1347]
080_EF1a T. brucei equiperdum	-----	-----	-----	-----	[774]
081_EF1a_Trypanosoma_avium	-----	-----	-----	-----	[1194]
082_EF1a_Trypanosoma_cruzi	-----	-----	-----	-----	[1347]
083_EF1a_Trypanosoma_rangeli	-----	-----	-----	-----	[1347]
084_EF1a_Ploeotia_costatum	-----	-----	-----	-----	[1229]
085_EF1a_Ploeotia_oblonga_cDNA	-----	-----	-----	-----	[1229]
086_EF1a_Ploeotia_oblonga_gDNA	-----	-----	-----	-----	[1167]
087_EF1a_Entosiphon_sp.	-----	-----	-----	-----	[1229]
088_EF1a_Entosiphon_sulcatum	-----	-----	-----	-----	[915]
089_EF1a_Peranema_tricophorum	-----	-----	-----	-----	[938]
090_EF1a P. tricophorum FJ807243	-----	-----	-----	-----	[1329]
091_EF1a_Colacium_elongatum	-----	-----	-----	-----	[1260]
092_EF1a_Colacium_mucronatum	-----	-----	-----	-----	[1103]
093_EF1a_Colacium_sp.	-----	-----	-----	-----	[1222]
094_EF1a_Colacium_vesiculosum	-----	-----	-----	-----	[1190]
095_EF1a_Cryptoglena_pigra	-----	-----	-----	-----	[1229]
096_EF1a_Cryptoglena_skujai	-----	-----	-----	-----	[1221]
097_EF1a_Discoplastis_adunca	-----	-----	-----	-----	[1229]
098_EF1a_Euglena_clara	-----	-----	-----	-----	[1156]
099_EF1a_Euglena_deses	-----	-----	-----	-----	[1229]
100_EF1a_Euglena_exilis	-----	-----	-----	-----	[1156]
101_EF1a_Euglena_geniculata	-----	-----	-----	-----	[1229]
102_EF1a_Euglena_gracilis	-----	-----	-----	-----	[1229]
103_EF1a E. gracilis X16890	-----	-----	-----	-----	[1335]
104_EF1a E. gracilis FJ807238	-----	-----	-----	-----	[1338]
105_EF1a_Euglena_granulata	-----	-----	-----	-----	[1229]
106_EF1a_Euglena_hiemalis	-----	-----	-----	-----	[1229]
107_EF1a_Euglena_longa	-----	-----	-----	-----	[852]
108_EF1a_Euglena_stellata	-----	-----	-----	-----	[1156]
109_EF1a_Euglena_viridis	-----	-----	-----	-----	[1172]
110_EF1a_Eutreptia_pertyi	-----	-----	-----	-----	[1229]
111_EF1a_Eutreptiella_pomquetensis	-----	-----	-----	-----	[839]
112_EF1a_Lepocinclis_acus_var._major	-----	-----	-----	-----	[1223]
113_EF1a_Lepocinclis_ovum	-----	-----	-----	-----	[1156]
114_EF1a_Lepocinclis_spyrogyroides_clone_38	-----	-----	-----	-----	[541]
115_EF1a L. spyrogyroides clones_1_&_4	-----	-----	-----	-----	[1227]
116_EF1a_Lepocinclis_steinii	-----	-----	-----	-----	[1156]
117_EF1a_Lepocinclis_tripteris	-----	-----	-----	-----	[1229]
118_EF1a_Monomorphina_aenigmatica	-----	-----	-----	-----	[1229]
119_EF1a_Monomorphina_pyrum	-----	-----	-----	-----	[1229]
120_EF1a_Monomorphina_sp.	-----	-----	-----	-----	[1228]
121_EF1a_Phacus_orbicularis	-----	-----	-----	-----	[1147]
122_EF1a_Phacus_similis	-----	-----	-----	-----	[1223]
123_EF1a_Strombomonas_acuminata	-----	-----	-----	-----	[1338]
124_EF1a_Strombomonas_borysteniensis	-----	-----	-----	-----	[1156]
125_EF1a_Strombomonas_napiformis	-----	-----	-----	-----	[1229]
126_EF1a_Strombomonas_ovalis	-----	-----	-----	-----	[1276]
127_EF1a_Strombomonas_triquetra	-----	-----	-----	-----	[891]
128_EF1a_Trachelomonas_conradi	-----	-----	-----	-----	[1191]
129_EF1a_Trachelomonas_echinata	-----	-----	-----	-----	[1226]
130_EF1a_Trachelomonas_lefevrei	-----	-----	-----	-----	[1223]
131_EF1a_Trachelomonas_planctonica	-----	-----	-----	-----	[1156]
132_EF1a_Trachelomonas_similis	-----	-----	-----	-----	[1156]

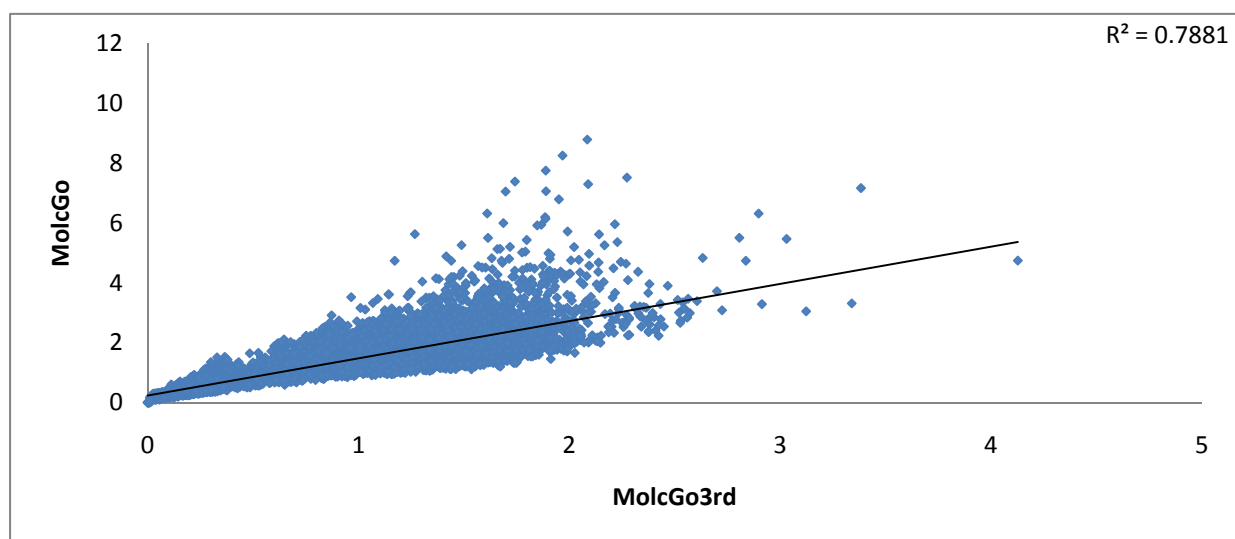
	1770	1780	1790	1800	
133_EFL1a_Trachelomonas_volvocina	-----	-----	-----	-----	[1223]
134_EFL environmental_sample_EU810232	-----	-----	-----	-----	[1389]
135_EFL_Creolimax_fragrantissima	-----	-----	-----	-----	[994]
136_EFL_Sphaeroforma_arctica	-----	-----	-----	-----	[1255]
137_EFL_Monosiga_brevicollis	-----	-----	-----	-----	[1386]
138_EFL_Xenopus_laevis	-----	-----	-----	-----	[1386]
139_EFL_Spizellomyces_punctatus	-----	-----	-----	-----	[1395]
140_EFL_Gaertneriomyces_semiglobiferous	-----	-----	-----	-----	[506]
141_EFL_Allomyces_macrognus	-----	-----	-----	-----	[1392]
142_EFL_Blastocladiella_emersonii	-----	-----	-----	-----	[1392]
143_EFL_Catenophylctis_sp.	-----	-----	-----	-----	[506]
144_EFL_Coelomomyces_stegomyiae	-----	-----	-----	-----	[506]
145_EFL_Entomophthora_muscae	-----	-----	-----	-----	[503]
146_EFL_Basidiobolus_ranarum	-----	-----	-----	-----	[503]
147_EFL_Conidiobolus_coronatus	-----	-----	-----	-----	[503]
148_EFL_Tetraselmis_tetrathele	-----	-----	-----	-----	[1338]
149_EFL_Mesostigma_viride	-----	-----	-----	-----	[879]
150_EFL_Ostreococcus_tauri	-----	-----	-----	-----	[1401]
151_EFL_Nephroselmis_olivacea	-----	-----	-----	-----	[957]
152_EFL_Micromonas_pusilla	-----	-----	-----	-----	[1353]
153_EFL_Parachlorella_kessleri	-----	-----	-----	-----	[942]
154_EFL_Helicosporidium_sp.	-----	-----	-----	-----	[1407]
155_EFL_Gonium_pectorale	-----	-----	-----	-----	[1389]
156_EFL_Chlamydomonas_reinhardtii	-----	-----	-----	-----	[1389]
157_EFL_Scenedesmus_obliquus	-----	-----	-----	-----	[1389]
158_EFL "Oryza sativa"	-----	-----	-----	-----	[1389]
159_EFL_Chlorococcum_sp.	-----	-----	-----	-----	[1338]
160_EFL_Ochlochaete_hystrix	-----	-----	-----	-----	[1269]
161_EFL_Ulva_intestinalis	-----	-----	-----	-----	[1338]
162_EFL_Urospora_sp.	-----	-----	-----	-----	[1086]
163_EFL_Gymnochlorella_stellata_ppc	CAAGAAGAAG	GACGACAAGA	AAAAGTCTTC	GAAAAAGAAG AAA	[1530]
164_EFL_Lotharella_amoebiformis_ppc	CAAGAAAAAG	TCGGCAAGA	AGAAGAAA--	-----	[1266]
165_EFL_Chlorarachnion_reptans_ppc	AAAGAAAAAGT	ACAAAGAAAA	AGAAG----	-----	[1293]
166_EFL_Chondrus_crispus	-----	-----	-----	-----	[1337]
167_EFL_Gracilaria_changii	-----	-----	-----	-----	[1390]
168_EFL_Cryptomonas_ovata	-----	-----	-----	-----	[1305]
169_EFL_Guillardia_theta	-----	-----	-----	-----	[1416]
170_EFL_Rhodomonas_salina	-----	-----	-----	-----	[1410]
171_EFL_Pavlova_lutheri	-----	-----	-----	-----	[1206]
172_EFL_Chrysochromulina_sp.	-----	-----	-----	-----	[1017]
173_EFL_Isochrysis_galbana	-----	-----	-----	-----	[1428]
174_EFL_Emilia_huxleyi	-----	-----	-----	-----	[1431]
175_EFL_Achnanthes_kuwaitensis	-----	-----	-----	-----	[1302]
176_EFL_Asterionella_glacialis	-----	-----	-----	-----	[1293]
177_EFL_Detionula_confervacea	-----	-----	-----	-----	[1299]
178_EFL_Thalassiosira_pseudonana	-----	-----	-----	-----	[1359]
179_EFL_Perkinsus_marinus	-----	-----	-----	-----	[1413]
180_EFL_Oxyrrhis_marina	-----	-----	-----	-----	[1440]
181_EFL_Karlodinium_micrum	-----	-----	-----	-----	[1425]
182_EFL_Amphidinium_carterae	-----	-----	-----	-----	[1351]
183_EFL_Heterocapsa_triquetra	-----	-----	-----	-----	[1422]
184_EFL_Planoglabratella_operculis	-----	-----	-----	-----	[961]
185_EFL_Reticulomyxa_filosa	-----	-----	-----	-----	[1332]
186_EFL_Lotharella_amoebiformis	-----	-----	-----	-----	[1506]
187_EFL_Gymnochlorella_stellata	-----	-----	-----	-----	[1488]
188_EFL_Chlorarachnion_reptans	-----	-----	-----	-----	[1419]
189_EFL_Diplonema_ambulator	-----	-----	-----	-----	[1011]
190_EFL_Diplonema_papillatum	-----	-----	-----	-----	[1392]
191_EFL_Rhynchopus_euleeides	-----	-----	-----	-----	[1032]
192_EFL_Petalomonas_cantuscygni	-----	-----	-----	-----	[924]
193_EFL_Trypanoplasma_borrelia	-----	-----	-----	-----	[1017]
194_EFL_Neobodo_salienis	-----	-----	-----	-----	[1278]

Figure 1.2. Saturation plots of nucleotide alignments. Alignments with all codon positions included (y-axis) vs. their respective alignment with the third codon positions removed (x-axis). Pairwise ML distances were calculated within each data set individually in PAUP* based on empirical values directly from the data or from empirical estimations of invariant sites and gamma shape parameters as reported by RAxML. **A.** Mo (1803 characters) vs. Mo3rd (1389 characters), **B.** MolcGo (1248 characters) vs. MolcGo3rd (1202 characters), **C.** MoGo (926 characters) vs. MoGo3rd (832 characters).

A.



B.



C.

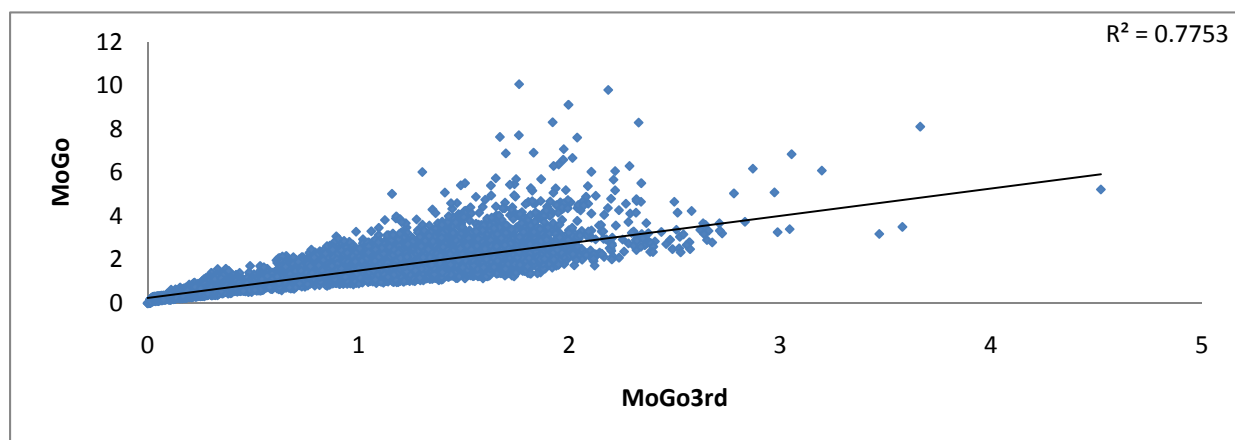


Figure 1.3. The Mo amino acid alignment of 194 OTUs. The OTUs are grouped according first to gene, then to organisms that were related in previous literature and are given a number from 001-194 which is consistent throughout the pages of the alignment. The column positions, from 1-601, are given in twelve column increments at the top of each page. Questionably named, identified or contaminated sequences are in quotation marks. The MolcGo alignment excludes the following columns numbers, with blocks excluded in brackets: [1-18], [184-191], [212], [220-239], [241-246], [269-272], [277-278], [290-295], [331], [369-373], [418-420], [422], [425], [430-433], [445], [453-463], [494-505], [520], [556-601]. The MoGo alignment excludes all columns that the MolcGo alignment does as well as the following additional columns, with blocks excluded in brackets: [144-152], [183], [213-219], [240], [268], [309-311], [365-364], [404], [421], [423-424], [426-429], [434-435], [451-452], [484-488], [519].

	10	20	30	40	50	60	70	80	90	100	110	
001_aEF1a_Desulfurococcus_mobilis	-----	--MSAPQ-KP	HLNIVIIGHV	DHGKSTMTGH	ILYRLGYFDE	KTVKMIIEES	KKMGKESFKF	AWLLDRMKEE	RERGVTSISL	YMKFETKKYF	FTIIDAPGHR	[97]
002_aEF1a_Sulfolobus_solfataricus	-----	---MSQ-KP	HLNLIVIGHI	DHGKSTLVGR	LLMDRGFIDE	KTVKAEAEAA	KKLGKSEKFF	AFLLDRLKEE	RERGVNTINL	FMRFETKKYF	FTIIDAPGHR	[95]
003_aEF1a_Thermoplasma_acidophilum	-----	---MKIMASQ-KP	HLNLITIGHV	DHGKSTLVGR	LLYEHGEIPA	HIIEEYRKEA	EQKGKATFEF	AWVMDRFKEE	RERGVITDLA	HRKFETDKYY	FTIIDAPGHR	[99]
004_aEF1a_Halobacterium_sp.	-----	---MSD-NR	HQNLAVIGHV	DHGKSTMVGR	LLYETGSVPE	HVIEQHKEEA	EEKGKGGSFF	AYVMDNLAAE	RERGVITIDIA	HQFETTDYEY	FTIVDCPGHR	[95]
005_aEF1a_Methanocaldococcus_jannaschii	-----	---MAKQ-KP	HLNVAFIGHV	DAGKSTTVGR	LLYDSGAIDP	QLELELKREA	ERKGKAGFEF	AYVMDNLKEE	RERGVITIDIA	HKFETTKQYV	FTIVDCPGHR	[96]
006_EF1a_Acanthamoeba_culbertsoni	-----	-----	---ELVIGHV	DAGKSTTTGH	LIYKCGGIDK	RTIEKFETE	KEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETAKYY	FTIIDAPGHR	[87]
007_EF1a_Neoparamoeba_branchipila	-----	-----	-----	DAGKSTTTGH	LIYKCGGIDK	RAIEKFEKEA	ADMKGGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETAKFY	FTIIDAPGHR	[80]
008_EF1a_Planoprotostelium_aurantium	-----	-----	-----	-AGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	KEIGKASFYK	AWVLDKLKAE	RERGITIDIA	LWKFETTKYY	FTIIDAPGHR	[79]
009_EF1a_Physarum_polycephalum	-----	-----	-----	-AGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKGSFKY	AWVLDKLKSE	RERGITIDIA	LWKFETAKYY	FTIIDAPGHR	[79]
010_EF1a_Amoebidium_parasiticum	-----	-----	---LVIGHV	DSGKSTSTGH	LIYKCGGIDK	RTIEKFEKEA	QEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFESTKYQ	FTIIDAPGHR	[86]
011_EF1a_Capsaspora_owczaraki	-----	---MGKE-KT	HINIVVIGHV	DSGKSTSTGH	LIYKCGGIDK	RTIEKFEKEA	QEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETSKFY	FTIIDAPGHR	[96]
012_EF1a_Monosiga_ovata	-----	-----	---GSVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETTKFY	FTIIDAPGHR	[87]
013_EF1a_Salpingoeca_amphoridium	-----	-----	---VIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKFY	FTIIDAPGHR	[85]
014_EF1a_Ceanorhabditis_elegans	-----	---MGKE-KV	HINIVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	QEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETAKYY	FTIIDAPGHR	[96]
015_EF1a_Drosophila_melanogaster	-----	---MGKE-KI	HINIVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	QEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETAKYY	FTIIDAPGHR	[96]
016_EF1a_Ornithorhynchus_anatinus	-----	---MGKE-KT	HINIVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIS	LWKFETTKYY	FTIIDAPGHR	[96]
017_EF1a_Xenopus_laevis	-----	---MGKE-KI	HINIVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIS	LWKFETGKYY	FTIIDAPGHR	[96]
018_EF1a_Saccharomyces_cerevisiae	-----	---MGKE-KS	HINVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AELGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYQ	FTIIDAPGHR	[96]
019_EF1a_Neurospora_crassa	-----	---MGKEDKT	HINVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AELGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYY	FTIIDAPGHR	[97]
020_EF1a_Cryptococcus_neoformans	-----	---MGKD-KL	HVNVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	QELGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPRYQ	FTIIDAPGHR	[96]
021_EF1a_Schizophyllum_commune	-----	---MGKE-KL	HVNVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AELGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYM	FTIIDAPGHR	[96]
022_EF1a_Monoblepharis_insignis	-----	-----	-----	-----	-----	---TEKFEKEA	AELGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWRFETPKYV	FTIIDAPGHR	[58]
023_EF1a_Rhizophlyctis_harderi	-----	-----	-----	-----	-----	---TEKFEKEA	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYY	FTIIDAPGHR	[58]
024_EF1a_Harpochytrium_sp.	-----	-----	-----	-----	-----	-----	---KY	AWVLDKLKAE	RERGITIDIA	LWKFETPKLY	FTIIDAPGHR	[42]
025_EF1a_Absidia_glauca	-----	---MGKE-KT	HVNVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AELGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYH	FTIIDAPGHR	[96]
026_EF1a_Smittium_simulii	-----	---ELVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	NELGKGSFKY	AWVLDKLKAE	RERGITIDIS	LWKFETPKYL	FTIIDAPGHR	[87]	
027_EF1a_Chara_australis	-----	-----	---SGKSTTTGH	LIYKLGIDK	RVIERFEKEA	AEMNKRSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETTRYV	FTIIDAPGHR	[79]	
028_EF1a_Physcomitrella_patens	-----	---MGKE-KV	HINIVVIGHV	DSGKSTTTGH	LIYKLGIDK	RVIERFEKEA	AEMNKRSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETMKYY	CTVIDAPGHR	[96]
029_EF1a_Acetabularia_acetabulum	-----	-----	-----	---SGKSTTTGH	LIYKLGIDK	RVIEKFEKEA	AEMNKRSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYH	CTVIDAPGHR	[79]
030_EF1a_Chaetomorpha_coliformis	-----	-----	-----	-----	-----	-----	-----	---ERGITIDIA	LWKFETPKYY	CTVIDAPGHR	[29]	
031_EF1a_Arabidopsis_thaliana	-----	---MGKE-KF	HINIVVIGHV	DSGKSTTTGH	LIYKLGIDK	RVIERFEKEA	AEMNKRSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETTKYY	CTVIDAPGHR	[96]
032_EF1a_Oryza_sativa	-----	---MGKE-KT	HINIVVIGHV	DSGKSTTTGH	LIYKLGIDK	RVIERFEKEA	AEMNKRSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETTKYY	CTVIDAPGHR	[96]
033_EF1a_Cyanidioschyzon_merolae	-----	---MGKE-KT	HVSVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	NEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETDKYY	FTIIDAPGHR	[96]
034_EF1a_Porphyrha_yezeensis	-----	---MGKE-KQ	HVSIVVIGHV	DSGKSTTTGH	LIYKCGGIEK	RAIEKFEKEA	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETEKYS	FTIIDAPGHR	[96]
035_EF1a_Plasmodium_falciparum	-----	---MGKE-KT	HINLVVIGHV	DSGKSTTTGH	LIYKLGIDK	RTIEKFEKES	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPRYF	FTIIDAPGHR	[96]
036_EF1a_Cryptosporidium_parvum	-----	---MGKE-KT	HINLVVIGHV	DSGKSTTTGH	LIYKLGIDK	RTIEKFEKES	SEMKGGSFKY	AWVLDKLKAE	RERGITIDIA	LWQFETPKYH	FTIIDAPGHR	[96]
037_EF1a_Tetrahymena_pyriformis	-----	---MARGDKV	HINLVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RVIEKFEKES	AELGKGSFKY	AWVLDKLKAE	RERGITIDIS	LWKFETAKYH	FTIIDAPGHR	[97]
038_EF1a_Stentor_coeruleus	-----	-----	---V	DSGKSTSTGH	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFKY	AWVLDKLKAE	RERGITIDIS	LFKFQTEKFY	FTIIDAPGHR	[81]
039_EF1a_Phytophthora_parasitica	-----	---MGKE-KV	HISLVVIGHV	DAGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AELGKTSFKY	AWVLDNLKAE	RERGITIDIA	LWKFESPKYF	FTIIDAPGHR	[96]
040_EF1a_Mallomonas_rasilis	-----	-----	---VIGHV	DAGKSTTTGH	LIYKCGGIDK	RTIEKYEKEA	AEMGKSSFKY	AWVLDNLKAE	RERGITIDIA	LWKFESPKFQ	FTIIDAPGHR	[85]
041_EF1a_Phaeodactylum_tricornutum	-----	---MGKE-KV	HISLVVIGHV	DAGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AELGKGSFKY	AWVLDNLKAE	RERGITIDIA	LWKFESPKYS	FTIIDAPGHR	[96]
042_EF1a_Thalassiosira_pseudonana	-----	---MPKD-KT	HISLVVIGHV	DGKSTTTGH	LIYKLGIDK	RTIEKYEKEA	NELGKGSFKY	AWVLDRLKAE	RERGITIDIA	LQTFETSKFY	FTIIDAPGHR	[96]
043_EF1a_Malawimonas_jakobiformis	-----	-----	-----	---TGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDKLKAE	RERGITIDIA	LWKFETPKYV	FTIIDAPGHR	[73]
044_EF1a_Malawimonas_californiana	-----	---MGKE-KT	HVNVVIGHV	DVGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDKLKAE	RERGITIDIA	LWKFETPKYY	FTIIDAPGHR	[96]
045_EF1a_Trimastix_pyriformis	-----	-----	---TTGH	LIYKCGGIDE	RTIQKFAEAA	NAIGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETEKYY	FTIIDAPGHR	[74]	
046_EF1a_Dinenympha_exilis	-----	-----	---STTTGH	LIYKCGGIDE	RTIKKFEQES	EAMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETNKYY	FTIIDAPGHR	[76]	
047_EF1a_Monocercomonoides_sp.	-----	---MGKE-KA	HINLVVIGHV	DVGKSTTTGH	LIYKCGGIDK	RTIEKFEQEA	DQIGKASFYK	AWVLDKLKAE	RERGITIDIA	LWKFETNKYY	FTIIDAPGHR	[96]
048_EF1a_Streblospiostrix_strix	-----	-----	---VGKSTTTGH	LIYKCGGIDK	RTIEKFEQEA	EQIGKASFYK	AWVLDKLKAE	RERGITIDIA	LWKFETNKYY	FTIIDAPGHR	[79]	
049_EF1a_Histomonas_meleagridis	-----	---MGKE-KE	HINLVVIGHV	DAGKSTTTGH	LIYKCGGIDK	RKLAQIEKEA	VQLGKGSFKY	AFVMDNLKAE	RERGITIDIS	LWKFESPKYM	FTIIDAPGHR	[96]
050_EF1a_Trichomonas_vaginalis	-----	---MGKE-KE	HINIVVIGHV	DAGKSTTTGH	LIYKCGGLDK	RKLAQIEKEA	EQLGKSSFKY	AFVMDLSLAE	RERGITIDIS	LWKFEGQKFS	FTIIDAPGHR	[96]
051_EF1a_Hexamita_inflata	-----	-----	-----	---NGKSTLTGH	LIYKCGGIDQ	RTLDEYEKKA	NEIGKGSFKY	AWVLDQLKDE	RERGITINIA	LWKFETKKFI	FTIIDAPGHR	[79]
052_EF1a_Spiroplasma_vortens	-----	-----	-----	---NGKSTLTGH	LIYKCGGIDQ	RTLDEYEKKA	NELGKGSFKY	AWVLDQLKDE	RERGITINIA	LWKFETKKFI	FTIIDAPGHR	[79]
053_EF1a_Andalucia_incarcerata	-----	-----	-----	---SGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETKKYY	FTIIDAPGHR	[79]
054_EF1a_Seculamonas_ecuadoriensis	-----	---MGKE-KA	HLNLVIGHV	DSGKSTTTGH	LIYQCGGIDK	RTIEKFEKEA	NDIGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYV	CTIIDAPGHR	[96]
055_EF1a_Reclinomonas_americana	-----	---MGKE-KA	HLNLVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	NEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYV	CTIIDAPGHR	[96]
056_EF1a_Histiona_aroidea	-----	---MGKE-KA	HLNLVIGHV	DSGKSTTTGH	LIYQCGGIDK	RTIEKFEKEA	AEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYV	CTIIDAPGHR	[96]
057_EF1a_Jakoba_libera	-----	---MGKE-KL	HLNLVIGHV	DSGKSTTTGH	LIYQCGGIDK	RTIEKFEKEA	SEMKGGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYV	CTIIDAPGHR	[96]
058_EF1a_Stachyamoeba_lipophora	-----	---MGKE-KA	HINLVVIGHV	DAGKSTTTGH	LIYKLGIDK	RVIEKFEKEA	ADIGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETPKYV	FTIIDAPGHR	[96]
059_EF1a_Naegleria_gruberi	-----	-----	---TTGH	LIYKCGGIDK	RVIEKFEKEA	AEMGKSSFKY	AWVLDKLKAE	RERGITIDIA	LWKFESKKYV	FTIIDAPGHR	[74]	
060_EF1a_Naegleria_andersoni	-----	---AGKSTTTGH	LIYKCGGIDK	RVIEKFEKEA	AEMGKSSFKY	AWVLDKLKAE	RERGITIDIA	LWKFESKKYV	FTIIDAPGHR	[79]		
061_EF1a_Acrasis_rosea	-----	---SGKSTTTGH	LIYKCGGIDK	RVIEKFEKEA	ADMKGQSFKY	AWVMDKLKAE	RERGITIDIA	LRKFETSKTM	FTIIDAPGHR	[79]		
062_EF1a_Sawyeria_marylandensis	-----	---MVKD-KI	HINLVVIGHV	DAGKSTTTGH	LIYKCGGIDK	RVIEKFEKEA	QEMGKGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETKKYF	FTIIDAPGHR	[96]
063_EF1a_Bodo_saltans	-----	-----	---GKSTATGH	LIYKCGGIDK	RTIEKFEKEA	AEIGKASFYK	AWVLDKLKAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[78]	
064_EF1a_Rhynchobodo_sp.	-----	-----	---ATGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDKLKAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[74]	
065_EF1a_Rhynchomonas_nasuta	-----	---GKE-KV	HISLVVIGHV	DAGKSTATGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKSSFKY	AWVLDKLKAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[95]
066_EF1a_Dimastigella_trypaniformis	-----	-----	---V	DAGKSTATGH	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDKLKAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[81]

		10	20	30	40	50	60	70	80	90	100	110			
067	EF1a	Perkinsiella	amoebae	-----	-----	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKCV	FTIIDAPGHR	[71]		
068	EF1a	Leishmania	major	-----	---MGKD-KV	LIYKCGGIDK	RTIEKFEKEA	AEIGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[96]		
069	EF1a	Leishmania	infantum	-----	---MGKD-KV	LIYKCGGIDK	RTIEKFEKEA	AEIGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[96]		
070	EF1a	Leishmania	braziliensis	-----	---MGKD-KV	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[96]		
071	EF1a	Leishmania	tarentolae	-----	---MGKD-KV	LIYKCGGIDK	RTIEKFEKEA	AEIGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[96]		
072	EF1a	"L. tarentolae"	FJ807257	-----	-----	LIYKCGGIDK	RSIEKFEKEA	AQMKGASFYK	AWVLDRLKAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[71]		
073	EF1a	Leptomonas	costaricensis	-----	-----	-----	-----	-----	-----	-----SV	FTIIDAPGHR	[12]			
074	EF1a	Leptomonas	bifurcata	-----	-----	-----	-----	-----	-----	-----ESP	FTIIDAPGHR	[16]			
075	EF1a	Leptomonas	podlipaevi	-----	-----	-----	-----	-----	-----	-----SV	FTIIDAPGHR	[12]			
076	EF1a	Herpetomonas	muscarum	-----	-----	---QCGGIDK	RTIEKFEKEA	AEIGKASFYK	AWVLDLKLAE	RERGITIDIS	LWKFESPKSV	FTIIDAPGHR	[67]		
077	EF1a	Herpetomonas	pessoai	-----	-----	-----	-----	-----	-----	-----DIA	LWKFESPKSV	FTIIDAPGHR	[23]		
078	EF1a	Blastocrithidia	culicis	-----	-----	---QCGGIDK	RTIEKFEKEA	ADIGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[67]		
079	EF1a	Trypanosoma	brucei	-----	---MGKE-KV	LIYKCGGIDK	RTIEKFEKEA	ADIGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[96]		
080	EF1a	T. brucei	equiperdum	-----	-----	LIYKCGGIDK	RTIEKFEKEA	ADIGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[70]		
081	EF1a	Trypanosoma	avium	-----	-----	---GKSTATGH	LIYKCGGIDK	RTIEKFEKEA	AEIGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[78]	
082	EF1a	Trypanosoma	cruzi	-----	---MGKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEIGKSSFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[96]		
083	EF1a	Trypanosoma	rangeli	-----	---MGKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEIGKSSFYK	AWVLDLKLAE	RERGITIDIA	LWKFESPKSV	FTIIDAPGHR	[96]		
084	EF1a	Ploetia	costatum	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
085	EF1a	Ploetia	oblonga_cDNA	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
086	EF1a	Ploetia	oblonga_gDNA	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
087	EF1a	Entosiphon	sp.	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFETPKSV	FTIIDAPGHR	[95]		
088	EF1a	Entosiphon	sulcatum	-----	-----	---H	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFETPKSV	FTIIDAPGHR	[71]	
089	EF1a	Peranema	tricophorum	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
090	EF1a	P. tricophorum	FJ807243	-----	---D-KV	LIYKCGGIDK	RTIEKFEKEA	AEMGKASFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[93]		
091	EF1a	Colacium	elongatum	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
092	EF1a	Colacium	mucronatum	-----	-----	---KSTTTGH	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[77]	
093	EF1a	Colacium	sp.	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
094	EF1a	Colacium	vesiculosum	-----	-----	DSG?STTTGH	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[81]	
095	EF1a	Cryptoglena	pigra	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
096	EF1a	Cryptoglena	skujai	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
097	EF1a	Discoplastis	adunca	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
098	EF1a	Euglena	clara	-----	-----	-----FEKEA	?EIGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[55]			
099	EF1a	Euglena	deses	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
100	EF1a	Euglena	exilis	-----	-----	---FEKEA	AEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[55]			
101	EF1a	Euglena	geniculata	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
102	EF1a	Euglena	gracilis	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
103	EF1a	E. gracilis	X16890	-----	---MGKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[96]		
104	EF1a	E. gracilis	FJ807238	-----	---MGKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[96]		
105	EF1a	Euglena	granulata	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
106	EF1a	Euglena	hiemalis	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
107	EF1a	Euglena	longa	-----	-----	-----	-----	-----	-----	-----	-----	[0]			
108	EF1a	Euglena	stellata	-----	-----	-----FEKEA	?E?GKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[55]			
109	EF1a	Euglena	viridis	-----	-----	---V	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	AEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[81]
110	EF1a	Eutreptia	pertyi	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
111	EF1a	Eutreptiella	pomquetensis	-----	-----	-----	-----	-----	-----	-----	-----	[0]			
112	EF1a	Lepocinlis	acus_var._major	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	NDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
113	EF1a	Lepocinlis	ovum	-----	-----	-----FEKEA	AE?GKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[55]			
114	EF1a	Lepocinlis	spirogyroides_clone_38	-----	-----	-----	-----	-----	-----	-----	-----	[0]			
115	EF1a	L. spirogyroides	clones_1_&_4	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
116	EF1a	Lepocinlis	steinii	-----	-----	-----FEKEA	GEIGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[55]			
117	EF1a	Lepocinlis	tripteris	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	NDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
118	EF1a	Monomorphina	aenigmatica	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
119	EF1a	Monomorphina	pyrum	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
120	EF1a	Monomorphina	sp.	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
121	EF1a	Phacus	orbicularis	-----	-----	---V	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	TDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[81]
122	EF1a	Phacus	similis	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	NDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
123	EF1a	Strombomonas	acuminata	-----	---MGKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[96]		
124	EF1a	Strombomonas	borysteniensis	-----	-----	-----FEKEA	AEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[55]			
125	EF1a	Strombomonas	napiformis	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
126	EF1a	Strombomonas	ovalis	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
127	EF1a	Strombomonas	triquetra	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
128	EF1a	Trachelomonas	conradi	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
129	EF1a	Trachelomonas	echinata	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	TDMGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
130	EF1a	Trachelomonas	lefevrei	-----	---GKE-KV	LIYKCGGIDK	RTIEKFEKEA	AEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]		
131	EF1a	Trachelomonas	planctonica	-----	-----	-----FEKEA	GEMKGGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[55]			
132	EF1a	Trachelomonas	similis	-----	-----	-----FEKEA	?EIGKGSFYK	AWVLDLKLAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[55]			

		10	20	30	40	50	60	70	80	90	100	110			
133	EFL1a	Trachelomonas_volvocina	-----	---GKE-KV	H?SLVVIGHV	DSGKSTTTGH	LIYKCGGIDK	RTIEKFEKEA	SEMKGGSFKY	AWVLDKLKAE	RERGITIDIA	LWKFETAKSV	FTIIDAPGHR	[95]	
134	EFL	environmental_sample_EU810232	-----	---MTEG-KI	HLSIVICGHV	DAGKSTTTGR	LIFELGGIPE	REMTKLRAEA	EALGKGSFSAF	AFYMDRQKDE	RARGVTIQT	TKEFFTDSYH	YTIIDAPGHR	[96]	
135	EFL	Creolimax_fragrantissima	-----	-----	-----	-----	---GH	LIFALGGIPE	REMTKLKEEA	TRLGKGSFSAF	AFYMDRQKDE	RERGVTIAC	TKEFFPTPKH	YTIIDAPGHR	[72]
136	EFL	Sphaeroforma_arctica	-----	---MTDG-KQ	HVSIVICGHV	DAGKSTTTGH	LIFALGGIPE	REMQKLKEEA	TRLGKGSFSAF	AFYMDRQKDE	RERGVTIAC	TKEFFPTPKH	YTIIDAPGHR	[96]	
137	EFL	Monosiga_brevicollis	-----	---MTDG-KQ	HVSIVICGHV	DAGKSTTTGR	LIFELGGIPE	REMQKLKDEA	ERLKGGSFSAF	AFYMDRQKEE	RERGVTIAC	TKEFFATATKH	YTIIDAPGHR	[96]	
138	EFL	Xenopus_laevis	-----	---MSEG-KQ	HVSIVICGHV	DAGKSTTTGR	LIFELGGIPE	REMQKLKDEA	QKLKGGSFSAF	AFYMDRQKDE	RERGVTIAC	TKEFFTTTKH	YTIIDAPGHR	[96]	
139	EFL	Spizellomyces_punctatus	-----	---MSGE-KP	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	REMEKLQREA	EHLGKSSFAF	AFYMDRQKEE	RARGVTIAC	TKEFFTTNNYH	YTIIDAPGHR	[96]	
140	EFL	Gaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
141	EFL	Allomyces_macrogynus	-----	---MSQG-KP	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	REMEKLQREA	DALGKSSFAF	AFYMDRQKEE	RARGVTIAC	TKEFFTDNRH	YTIIDAPGHR	[96]	
142	EFL	Blastocladiella_emersonii	-----	---MSEG-KP	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	REMEKLRAEA	DALGKSSFAF	AFYMDRQKEE	RARGVTIAC	TKEFFTEKWH	YTIIDAPGHR	[96]	
143	EFL	Catenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
144	EFL	Coelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
145	EFL	Entomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
146	EFL	Basidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
147	EFL	Conidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
148	EFL	Tetraselmis_tetrathele	-----	-----	-----	---SGKSTTTGR	LLFELGGIPE	REMDKLKAEA	EALGKQSFSAF	AFYMDRQKEE	RERGVTISCT	TKEFFTDKYH	YTIIDAPGHR	[79]	
149	EFL	Mesostigma_viride	-----	---MAEANKT	HLSIVICGHV	DSGKSTTTGR	LIFELGGIPE	RDMKLRAEA	DRLGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFYTDKWH	YTIIDAPGHR	[97]	
150	EFL	Ostreococcus_tauri	-----	---MSEG-KE	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKAEA	DALGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFTEKWH	YTIIDAPGHR	[96]	
151	EFL	Nephrolepis_olivacea	-----	-----	---SIVICGHV	DSGKSTTTGR	LLFELGGIPE	REMEKLRAEA	DRLGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFTEKWH	YTIIDAPGHR	[88]	
152	EFL	Micromonas_pusilla	-----	-----	-----	---SGKSTTTGR	LIFELGGIPE	RETEKLRAEA	DALGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFTDNRH	YTIIDAPGHR	[79]	
153	EFL	Parachlorella_kessleri	-----	-----	---HV	DSGKSTTTGR	LLFELGGIPE	RELEKLKEEA	SALGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFTDNRH	YTIIDAPGHR	[82]	
154	EFL	Helicosporidium_sp.	-----	---MAEE-KP	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKEEA	AALGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFTKNWH	YTIIDAPGHR	[96]	
155	EFL	Gonium_pectorale	-----	---MSEG-KE	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKEEA	AALGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFATAKWH	YTIIDAPGHR	[96]	
156	EFL	Chlamydomonas_reinhardtii	-----	---MSEG-KE	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKEEA	AALGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFTDNRH	YTIIDAPGHR	[96]	
157	EFL	Scenedesmus_obliquus	-----	---MAEG-KQ	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKEEA	AALGKQSFSAF	AFYMDRQKEE	RERGVTIAC	TKEFFTEKWH	YTIIDAPGHR	[96]	
158	EFL	"Oryza_sativa"	-----	---MAEG-KE	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKEEA	AALGKQSFSAF	AFYMDRQKEE	RERGVTIAC	TKEFFTDKWH	YTIIDAPGHR	[96]	
159	EFL	Chlorococcum_sp.	-----	-----	-----	---SGKSTTTGR	LIFELGGISD	REVEKLKEEA	ATLGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFYTNNWH	YTIIDAPGHR	[79]	
160	EFL	Ochlochaeta_hystrix	-----	-----	---HV	DSGKSTTTGR	LLFELGGIPP	RELDKLKEEA	AALGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFYTNRWH	YTIIDAPGHR	[82]	
161	EFL	Ulva_intestinalis	-----	-----	---SGKSTTTGR	LIFELGGIPE	RELEKLKEEA	AHLGKASFAF	AFYMDRQKEE	RERGVTISCT	TKEFFTENNWH	YTIIDAPGHR	[79]		
162	EFL	Urospora_sp.	-----	-----	-----	-----	-----	-----	-----	-----	RERGVTISCT	TKEFFTDNRH	YTIIDAPGHR	[30]	
163	EFL	Gymnochlorella_stellata_ppc	---GGGDAAAA	PADSSEPLKP	HLSIVICGHV	DSGKSTTTGR	LLFELGGIEE	REMERLKSEA	KALNKETFLF	AFYMDRQKEE	RERGVTITCA	TKEFFTEKWH	YTIIDAPGHR	[108]	
164	EFL	Lotharella_amoebiformis_ppc	-----	-----	-----	-----	-----	-----	-----	---TCA	TKEFFTEKWH	YTIIDAPGHR	[23]		
165	EFL	Chlorarachnion_reptans_ppc	-----	-----	-----	-----	-----	-----	---EE	RERGVTITCA	TKEFFTKKWH	YTIIDAPGHR	[32]		
166	EFL	Chondrus_crispus	-----	---MTEG-KK	HVSIVICGHV	DAGKSTTTGR	LIFELGGISE	REMQKLKDEA	KALGKDSFSAF	AFYMDKSKEE	RARGVTIAC	TKEFFTDSYH	YTIIDAPGHR	[96]	
167	EFL	Gracilaria_changii	-----	---MTEG-KK	HVSIVICGHV	DAGKSTTTGR	LIFELGGISD	REMQKLKDEA	KALGKDSFSAF	AFYMDKSKEE	RARGVTIAC	TKEFFTDSYH	YTIIDAPGHR	[96]	
168	EFL	Cryptomonas_ovata	-----	-----	---V	DAGKSTTTGR	LLFELGGIPE	RELEKLKEEA	ANLGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFTEKWH	YTIIDAPGHR	[81]	
169	EFL	Guillardia_theta	-----	---MAE-KE	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKEEA	ANLGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFTDKWH	YTIIDAPGHR	[95]	
170	EFL	Rhodomonas_salina	-----	---MAE-KA	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKEEA	ANLGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFTDKWH	YTIIDAPGHR	[95]	
171	EFL	Pavlova_lutheri	-----	---MPE-KE	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	REMEKLRAEA	DRLGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFTEKWH	YTIIDAPGHR	[95]	
172	EFL	Chrysochromulina_sp.	-----	-----	-----	-----	-----	---RQKEE	RERGVTISCT	TKEFFTEKWH	YTIIDAPGHR	[35]			
173	EFL	Isochrysis_galbana	-----	---MASE-KD	HMSVVICGHV	DSGKSTTTGR	LLFELGGIPE	REMEKLKAEA	DRLGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFTEKWH	YTIIDAPGHR	[96]	
174	EFL	Emiliana_huxleyi	-----	---MAE-KD	HMSVVICGHV	DSGKSTTTGR	LLFEMGGIPE	REMEKLKAEA	DRLGKSSFAF	AFYMDRQKEE	RERGVTISCT	TKEFFTDKWH	YTIIDAPGHR	[95]	
175	EFL	Achnanthes_kuwaitensis	-----	-----	-----	---TCGH	LIFKQGGISE	REMTKLQAMA	EEKGKSSFGF	AYYLDTCKEE	RERGVTIQCN	TKEFFTDKYH	YTIIDAPGHR	[74]	
176	EFL	Asterionella_glacialis	-----	-----	-----	---TCGH	LIFKQGGISE	REMTKLQETA	EEKGKASFGF	AYYLDTCKEE	RERGVTIQCN	TKEFFTDKYH	YTIIDAPGHR	[74]	
177	EFL	Detonula_confervacea	-----	-----	---TCGH	LIFRQGGISA	REMEKLNETA	AEKGKSSFGF	AYYLDTCKEE	RERGVTIQCN	TKEFFTDKYH	YTIIDAPGHR	[74]		
178	EFL	Thalassiosira_pseudonana	-----	---MSG-KE	HLSVVVCGHV	DAGKSTTTGH	LIFKQGGISQ	REMDKLQATA	EEKGKSSFGF	AYYLDTCKEE	RERGVTIQCN	TKEFFTEKYH	YTIIDAPGHR	[95]	
179	EFL	Perkinsus_marinus	-----	---MSTGKEGKV	HMSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLQKEA	DRLGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFTDSWH	YTIIDAPGHR	[99]	
180	EFL	Oxyrrhis_marina	-----	---MGDKSEKE	HLSIVICGHV	DSGKSTTTGH	LLFELGGIPE	RELDKLQKEA	DRLGKSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFYTEKWH	YTIIDAPGHR	[98]	
181	EFL	Karlodinium_micrum	-----	---MTANSEKE	HLSIVICGHV	DSGKSTTTGR	LIFELGGIPE	RELEKLQKEA	ERLKGGSFSAF	AFYMDRQKEE	RERGVTIAC	TKEFFYTEKWH	YTIIDAPGHR	[98]	
182	EFL	Amphidinium_carterae	-----	---MPANSEKE	HLSIVICGHV	DSGKSTTTGR	LLFELGGIPE	RELEKLKAEA	ERLKGSSFAF	AFYMDRQKEE	RERGVTIAC	TKEFFYTEKWH	YTIIDAPGHR	[98]	
183	EFL	Heterocapsa_triquetra	-----	---MPANSEKE	HLSIVICGHV	DSGKSTTTGR	LIFELGGIPE	RELEKLQKEA	ERLKGGSFSAF	AFYMDRQKEE	RERGVTIAC	TKEFFYTEKWH	YTIIDAPGHR	[98]	
184	EFL	Planoglabratella_operculis	-----	-----	-----	-----	-----	---AF	AFYMDKSKEE	RARGVTIQCC	TREFFTDSYH	YTIIDAPGHR	[42]		
185	EFL	Reticulomyxa_filosa	-----	-----	-----	---R	LLFELGGMSD	RELEKLKKEA	QELGKDSFLF	AFYLDRSKEE	RARGVTIVCT	TREFFTPSWH	YTIIDAPGHR	[71]	
186	EFL	Lotharella_amoebiformis	MSAQKPANPS	ELNMKDPNKP	HVGVVITGHV	DAGKSTTTGH	LLFKLGTMD	REKRALIAEA	EKEGKGSFSAF	AFYMDKQKEE	RKRGVTISCT	TKEFFTTNNH	YTIIDAPGHR	[110]	
187	EFL	Gymnochlorella_stellata	-----	-----	---HV	DAGKSTTTGH	LLFELGTMD	RMKADLANKA	KEMKKSFAF	AFYMDKQKEE	RERGVTISCT	TKEFHTTNFH	YTIIDAPGHR	[82]	
188	EFL	Chlorarachnion_reptans	-----	-----	---H	DAGKSTTTGH	LLFELGTMD	RAKADLLAKA	KEMKKSFAF	AFYMDKQKEE	RERGVTISCT	TKEFHTTNFH	YTIIDAPGHR	[71]	
189	EFL	Diplonema_ambulator	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
190	EFL	Diplonema_papillatum	-----	---IVICGHV	DSGKSTTTGR	LLFELGGIPE	REMEKLQKEA	ERLKGQSFSAF	AFYMDRQKEE	RERGVTIAC	TKEFFTEKWH	YTIIDAPGHR	[87]		
191	EFL	Rhynchopus_euleeides	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
192	EFL	Petalomonas_cantuscyni	-----	-----	-----	-----	-----	---SFAF	LFYMDRCKEE	RERGVTIQT	LKEFFTPQFH	YTIIDAPGHR	[44]		
193	EFL	Trypanoplasma_borrelli	-----	---MAENEGKA	HVSIVVAGHV	DAGKSSLTGR	LLYDLGGIDE	RAEKLRLKIA	EENGKPSFAF	AYYMDNQKAE	RERGVTIACN	TKEFFYTERYH	YTIIDAPGHR	[98]	
194	EFL	Neobodo_saliens	-----	-----	---V	DAGKSSLTGR	LLYDMGGIDE	RTMEKLKKKA	EENGKPSFAF	AYYLDNQKAE	QERGVTIACN	TKEFFYTERYH	YTIIDAPGHR	[81]	

	120	130	140	150	160	170	180	190	200	210	220	
001_aEF1a_Desulfurococcus_mobilis	DFVKNMITGA	SQADAAILVV	SARKGEFEAG	MSAE-----	--GQTREHAI	LARTMGINDL	IVAINKMDAT	EPP-----	-YSEKRYNEI	KEILGKFLKG	L-GY-----	[184]
002_aEF1a_Sulfolobus_solfataricus	DFVKNMITGA	SQADAAILVV	SAKKGEYFAG	MSVE-----	--GQTREHII	LAKTMGLDQL	IVAVNKMMDLT	EPP-----	-YDEKRYKEI	VDQVSKFMRS	Y-GF-----	[182]
003_aEF1a_Thermoplasma_acidophilum	DFVKNMITGT	SQADAAILVI	SAREGEGVM-	-----	--EQTREHAF	LARTLGVQPM	VVAINKMDAT	SPP-----	-YSEKRYNEV	KADAELKLLRS	I-GF-----	[181]
004_aEF1a_Halobacterium_sp.	DFVKNMITGA	SQADANVLVV	AADGDGVA--	-----	--PQTREHVF	LSRTLGLIDEL	IVAVNKMMD--	VVD-----	-YDESKYNEV	VSVGKDLFGQ	V-GF-----	[173]
005_aEF1a_Methanocaldococcus_jannaschii	DFIKNMITGA	SQADAAILVV	DVNDAKTGIQ	-----	--PQTREHMF	LARTLGKIQI	AVAINKMDI--	TVN-----	-YSQEEYEM	KMLLSEQLLK	VLGY-----	[178]
006_EF1a_Acanthamoeba_culbertsoni	DFIKNMITGT	SQADAVILVI	ASGEFEFAG	ISKN-----	--GQTREHAL	LAFTLGVKQM	IVCNKMD--	NVN-----	-WAENRYNEI	QREVSGLYKK	V-GY-----	[172]
007_EF1a_Neoparamoeba_branchipila	DFIKNMITGT	SQADAVLVVI	ASGTGEFEAG	IAKN-----	--GQTREHAL	LSFTLGVKQM	ICAINKMDDK	SVN-----	-YSEDRYTEI	KKEVSGFLKK	I-GY-----	[167]
008_EF1a_Planoprotostelium_aurantium	DFIKNMITGT	SQADAILVI	ASPTGEFEAG	IDKN-----	--GQTREHAL	LAFTLGVKQL	IVLVNKMDEK	STN-----	-YSEARFNEI	RTEVSSFIKK	I-GY-----	[166]
009_EF1a_Physarum_polycephalum	DFIKNMITGT	SQADAAILVI	ASPTGEFEAG	IAKT-----	--GQTREHAL	LAYTLGVKQM	IVAINKMDEK	SVN-----	-WSQARYDEI	VKETSSFVKK	I-GY-----	[166]
010_EF1a_Amoebidium_parasiticum	DFIKNMITGT	SQADCAVLII	AAGTGEFEAG	ISKD-----	--GQTREHAL	LAFTLGVKQL	IVAINKMD--	SIK-----	-FAQDRFNEI	VNEVSNIYKK	I-GY-----	[171]
011_EF1a_Capsaspora_owczaraki	DFIKNMITGT	SQADCAVLIV	AAGTGEFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQL	IVGINKMD--	SIK-----	-FAEERYNEI	VTEVSNIYKK	I-GY-----	[181]
012_EF1a_Monosiga_ovata	DFIKNMITGT	SQADCAVLIV	AAGTGEFEAG	ISSN-----	--GQTREHAL	LAYTLGVKQL	IVGVNKMDS	EPP-----	-YSESRFNEI	KTEVSTYIKK	I-GY-----	[174]
013_EF1a_Salpingoeca_amphoridium	DFIKNMITGT	SQADCAVLIV	AAGTGEFEAG	ISSN-----	--GQTREHAL	LAYTLGVKQL	IVGVNKIDST	EPP-----	-YSEARFNEI	KTEVSNIYKK	V-GF-----	[172]
014_EF1a_Ceanorhabditis_elegans	DFIKNMITGT	SQADCAVLVV	ACGTGEFEAG	ISKN-----	--GQTREHAL	LAQTLGVKQL	IVACNKMDS	EPP-----	-FSEARFTEI	TNEVSGFIKK	I-GY-----	[183]
015_EF1a_Drosophila_melanogaster	DFIKNMITGT	SQADCAVLIV	AAGTGEFEAG	ISKN-----	--GQTREHAL	LAYTLGVKQL	IVGVNKMDS	EPP-----	-YSEARYEEI	KKEVSSYIKK	I-GY-----	[183]
016_EF1a_Ornithorhynchus_anatinus	DFIKNMITGT	SQADCAVLIV	AGVGGEFEAG	ISKN-----	--GQTREHAL	LAYTLGVKQL	IVGVNKMDS	EPP-----	-YSEKRYDEI	VKEVSAYIKK	I-GY-----	[183]
017_EF1a_Xenopus_laevis	DFIKNMITGT	SQADCAVLIV	AGVGGEFEAG	ISKN-----	--GQTREHAL	LAYTLGVKQL	IVGVNKMDS	EPP-----	-FSQKRFEEI	TKEVSAYIKK	I-GY-----	[183]
018_EF1a_Saccharomyces_cerevisiae	DFIKNMITGT	SQADCAILII	AGVGGEFEAG	ISKD-----	--GQTREHAL	LAYTLGVRQL	IVAVNKMMD--	SVK-----	-WDSERFQEI	VKETSNIYKK	V-GY-----	[181]
019_EF1a_Neurospora_crassa	DFIKNMITGT	SQADCAILII	AAGTGEFEAG	ISKD-----	--GQTREHAL	LAYTLGVRQL	IVAINKMD--	TTQ-----	-WSQTRFEEI	IKETKNFIKK	V-GY-----	[182]
020_EF1a_Cryptococcus_neoformans	DFIKNMITGT	SQADCAILII	ATGIGFEFAG	ISKD-----	--GQTREHAL	LAYTLGVRQL	IVAVNKMMD--	TCK-----	-WSEDRFNEI	VKETNGFIKK	V-GY-----	[181]
021_EF1a_Schizophyllum_commune	DFIKNMITGT	SQADCAILII	AGGTGEFEAG	ISKD-----	--GQTREHAL	LAYTLGVRQL	IVAVNKMMD--	TTK-----	-WSEDRFNEI	VKETSTFYKK	V-GY-----	[181]
022_EF1a_Monoblepharis_insignis	DFIKNMITGT	SQADCAILII	AGGTGEFEAG	ISKD-----	--GQTREHAL	LSFTLGVRQL	IVAVNKMMD--	STG-----	-YSEERFNEI	VREVSNIYKK	V-GY-----	[143]
023_EF1a_Rhizophlyctis_harderi	DFIKNMITGT	SQADCAILVI	ASGVGEFEAG	ISKD-----	--GQTREHAL	LANTLGVKQM	IVAVNKMMD--	TNK-----	-WSEARFNEI	VKETSSFYKK	V-GY-----	[143]
024_EF1a_Harpochytrium_sp.	DFIKNMITGT	SQADCAVLIV	PAGTGEFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVAVNKMMD--	SAK-----	-WSEERYNEI	VKEVSSYIKK	V-GY-----	[127]
025_EF1a_Absidia_glauca	DFIKNMITGT	SQADCGILII	AAGTGEFEAG	ISKD-----	--GQTREHAL	LAYTLGVRQL	IVAINKMD--	STK-----	-WSEQRFNEI	IKESVGFYKK	I-GF-----	[181]
026_EF1a_Smittium_simulii	DFIKNMITGT	SQADCAILII	AGVGGEFEAG	ISKD-----	--GQTREHAL	LAYTLGVRQL	IVAVNKMMD--	SNK-----	-YSEERFTEI	IKESVNIYKK	V-GY-----	[172]
027_EF1a_Chara_australis	DFIKNMITGT	SQADCAVLVI	DSTTGGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	ICACNKMDS	TPK-----	-YSENRNYEI	KKEVSTYIKK	V-GY-----	[166]
028_EF1a_Physcomitrella_patens	DFIKNMITGT	SQADCAVLII	DSTTGGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	ICCNCKMDS	TPK-----	-YSKARYEEI	KKEVSAYLKK	V-GY-----	[183]
029_EF1a_Acetabularia_acetabulum	DFIKNMITGT	SQADCAVLMI	DSTPGGFESG	TSKD-----	--GQTREHGL	LAYTLGVKQM	IVGVNKMDS	EPP-----	-YSDARYNEI	KTEVSYLYKK	V-GY-----	[166]
030_EF1a_Chaetomorpha_coliformis	DFIKNMITGT	TQADCAILMI	DSTPGGFESG	ISKD-----	--GQTREHAL	LAYTLGVKQM	ICCNCKMDS	EPP-----	-YSDKRYEEI	MKEVKNYIKK	V-GY-----	[116]
031_EF1a_Arabidopsis_thaliana	DFIKNMITGT	SQADCAVLII	DSTTGGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	ICCNCKMDS	TPK-----	-YSKARYDEI	IKESVSYLYKK	V-GY-----	[183]
032_EF1a_Oryza_sativa	DFIKNMITGT	SQADCAVLII	DSTTGGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	ICCNCKMDS	TPK-----	-YSKARYDEI	IKESVSYLYKK	V-GY-----	[183]
033_EF1a_Cyanidioschyzon_merolae	DFIKNMITGT	SQADLAILVV	ASPPGEFEAG	ISSN-----	--GQTREHAL	LAYTLGVKQM	IVAVNKMDDK	NVN-----	-WSQSRYDEI	TKELNLYLKK	V-GY-----	[183]
034_EF1a_Porphyrha_jezoensis	DFIKNMITGT	SQADLAILVI	ASPPGEFEAG	ISQN-----	--GQTREHAL	LAYTLGVKQM	IVAVNKMDDK	NVN-----	-WSQDRYEEV	SKEMDLYLKK	V-GY-----	[183]
035_EF1a_Plasmodium_falciparum	DFIKNMITGT	SQADVALLV	PAEVGGFEGA	FSKE-----	--GQTKHEAL	LAYTLGVKQM	IVGVNKMDDK	TVK-----	-YSEDRYEEI	SKEMDLYLKK	V-GY-----	[181]
036_EF1a_Cryptosporidium_parvum	DFIKNMITGT	SQADVALLV	PA--DRFEGA	FSKE-----	--GQTKHEAL	LAYTLGVRQM	IVGVNKMDDK	TCE-----	-YKQSRFDEI	FNEVDGYLKK	V-GY-----	[179]
037_EF1a_Tetrahymena_pyriformis	DFIKNMITGT	SQADVALMI	ASPPGEFEAG	ISKD-----	--GQTKHEAL	LAYTLGVKQM	IVGVNKMDDK	TVN-----	-FSEERYQEI	KNEVSYLYKK	V-GY-----	[184]
038_EF1a_Stentor_coeruleus	DFIKNMITGT	SQADVALIV	ASGKGEFEAG	IGKN-----	--GQTKHEAL	LAYTLGVKQM	IVGVNKMDDK	SCN-----	-WSQARYDEI	KNEVSQYLYKK	V-GY-----	[168]
039_EF1a_Phytophthora_parasitica	DFIKNMITGT	SQADCAILVV	ASGVGEFEAG	ISKE-----	--GQTKHEAL	LAYTLGVKQM	IVAVNKMDDK	SVN-----	-YGQARYEEI	KNEVSTYLYKK	V-GY-----	[183]
040_EF1a_Mallomonas_rasilis	DFIKNMITGT	SQADVALVI	DGAPGGFEGG	WGSD-----	--GQTKHEAL	LAYTLGVNQM	IVAVNKMDDK	GCN-----	-YSEQRYNDI	KNEVSTYLYKK	V-GY-----	[170]
041_EF1a_Phaeodactylum_tricornutum	DFIKNMITGT	SQADVALVI	DSSQGGFEAG	ISKD-----	--GQTKHEAL	LAYTLGVKQM	IVAVNKMDDK	TVK-----	-YAEADRYTEI	KNEVSAYLKK	V-GY-----	[183]
042_EF1a_Thalassiosira_pseudonana	DFIKNMITGT	SQADVALVI	DAAGKGFAG	ISKE-----	--GQTKHEAL	LAHTLGVRQV	MVVNKMDDK	SVQ-----	-YSEDRFNEI	KDEVSRYLTH	L-GY-----	[183]
043_EF1a_Malawimonas_jakobiformis	DFIKNMITGT	SQADAAILVV	AATTGEFEAG	ISKD-----	--GQTKHEAL	LAYTLGVKQI	VVAVNKMDDK	TVN-----	-FSQDRFEEI	VKEVSSFLKK	T-GY-----	[160]
044_EF1a_Malawimonas_californiana	DFIKNMITGT	SQADAAILVV	AAGTGEFEAG	ISKD-----	--GQTKHEAL	LAYTLGVKQM	VVAVNKMDDK	TVN-----	-WGQDRYEEI	KKEVSNIYKK	I-GY-----	[183]
045_EF1a_Trimastix_pyriformis	DFIKNMITGT	SQADCAILLV	AATAGEFEAG	ISGE-----	--GQTKHEAL	LAYTLGVKQM	IVLVNKMDDK	NQDPKGTGENK	PFLGARFDEI	KDNVFIKLKK	I-GY-----	[169]
046_EF1a_Dinenympha_exilis	DFIKNMITGT	SQADAAILVV	AANVGFEFAG	ISKD-----	--GQTKHEAL	LAYTLGVKQM	IVLVNKMDDK	SVN-----	-WAESRYNEI	KTEMRTYLYKK	I-GY-----	[163]
047_EF1a_Monocercomonoides_sp.	DFIKNMITGT	SQADVALLV	AANVGFEFAG	ISKD-----	--GQTKHEAL	LAYTLGVKQM	IVLVNKMDDK	SVN-----	-YSEARFNEI	KGEMRNYLYKK	I-GY-----	[183]
048_EF1a_Streblospio_trix	DFIKNMITGT	SQADAAILVV	AANQGEFEAG	ISKD-----	--GQTKHEAL	LAYTLGVKQM	IVLVNKMDDK	SVS-----	-FSEARYNEI	VGEMRNYLYKK	I-GY-----	[166]
049_EF1a_Histomonas_meleagridis	DFIKNMITGT	SQADAAILVI	DATRGGEFAG	IAEQ-----	--GQTKHEAL	LAYTLGKIQV	IVGVNKMDDN	TVN-----	-YSKQRYDEI	VQEMTHILGN	I-GF-----	[183]
050_EF1a_Trichomonas_vaginalis	DFIKNMITGT	SQADAAILVI	DSTLGGEFAG	IAEQ-----	--GQTKHEAL	LAYTLGKIQV	IVGVNKMDDK	TVN-----	-YKNARFDEI	TAEMTRITLG	I-GY-----	[183]
051_EF1a_Hexamita_inflata	DFIKNMITGT	SQADVALVV	AAGQGEFEAG	ISSE-----	--GQTKHEAT	LANTLGKIKM	IVAVNKMDDK	QVN-----	-YSEARYTEI	KTEMQKTFQ	I-GY-----	[166]
052_EF1a_Spiroplasma_vortens	DFIKNMITGT	SQADVALVV	AAGQGEFEAG	ISKE-----	--GQTKHEAT	LANTLGKIKM	IVAVNKMDDK	NVN-----	-YSKDRYNEI	KTEMQKTLVA	I-GY-----	[166]
053_EF1a_Amalucina_incarcerata	DFIKNMITGT	SQADAAILVV	ASGTGEFEAG	ISKE-----	--GQTKHEAL	LAYTLGVKQM	IVAVNKMDDK	SVN-----	-YGQARFEEI	KTEVSAYLKK	V-GY-----	[166]
054_EF1a_Seculamonas_ecuadoriensis	DFIKNMITGT	SQADAAILVI	ASGTGEFEAG	ISSE-----	--GQTKHEAL	LAYTLGVKQI	AVAINKIDDK	SVN-----	-YSQARYDEI	KAEVSAYLKK	V-GY-----	[183]
055_EF1a_Reclinomonas_americana	DFIKNMITGT	SQADAAILVI	ASGTGEFEAG	ISKE-----	--GQTKHEAL	LAYTLGVKQI	IVAVNKMDDK	SVN-----	-YSEARYNEI	KAEVSAYLKK	V-GY-----	[183]
056_EF1a_Histiona_roides	DFIKNMITGT	SQADAAILVV	ASGTGEFEAG	ISKE-----	--GQTKHEAL	LAYTLGVKQI	IVAVNKMDDK	SVN-----	-YGQDRYDEI	KAEVSAYLKK	V-GY-----	[183]
057_EF1a_Jakoba_libera	DFIKNMITGT	SQADAAILVV	ASGTGEFEAG	ISKE-----	--GQTKHEAL	LAYTLGVKQM	IVGVNKMDDK	SVN-----	-YSEARYTEI	KKEVGAYLKK	V-GY-----	[183]
058_EF1a_Stachyamoeba_lipophora	DFIKNMITGT	SQADVALVI	DSTGGEFAG	FARD-----	--GQTKHEAL	LAYTLGVKQF	VVAVNKMDDK	SIN-----	-YSQDRYEEI	KKNVADYLYKK	V-GY-----	[183]
059_EF1a_Naegleria_gruberi	DFIKNMITGT	SQADVALVV	DSTGGEFAG	FSKD-----	--GQTKHEAL	LAYTLGKIQM	IVCNKMFDDT	SVN-----	-YAEKRYDEI	KSEVGRYLYKK	L-GF-----	[161]
060_EF1a_Naegleria_andersoni	DFIKNMITGT	SQADVALVV	DSTGGEFAG	FSKD-----	--GQTKHEAL	LAYTLGKIQM	IVCNKMFDDT	SVS-----	-YKEDRYNEI	KSEVGRYLYKK	L-GF-----	[166]
061_EF1a_Acrasis_rosea	DFIKNMITGT	SQADAAILVV	DSTGGEFAG	ISKD-----	--GQTKHEAL	LAYTLGVKQM	IVCNKMFDDT	SVQ-----	-YKEDRYKEI	KNEVADYLYKK	V-GY-----	[166]
062_EF1a_Sawyeria_marylandensis	DFIKNMITGT	SQADAAMLVI	DSAKGGEFAG	IAEE-----	--GQTKHEAL	LAYTLGVKQM	IVGVNKMDDK	TTM-----	-YSQERYTEI	QKGVSDIYKK	V-GY-----	[183]
063_EF1a_Bodo_saltans	DFIKNMITGT	SQADAAILVI	ASGTGEFEAG	FSKD-----	--GQTKHEAL	LAYTLGVKQM	VVAVNKMDDK	SIN-----	-FSQARYDEI	VSNVGQYLYKK	V-GY-----	[163]
064_EF1a_Rhynchobodo_sp.	DFIKNMITGT	SQADAAILVI	ASGQGEFEAG	ISKE-----	--GQTKHEAL	LAYTLGVKQM	IVAVNKMDDK	SVN-----	-YSEARYTEI	MKEVSAYLKK	V-GY-----	[161]
065_EF1a_Rhynchomonas_nasuta	DFIKNMITGT	SQADAAILVI	ASGQGEFEAG	YAKE-----	--GQTKHEAL	LAYTLGVKQM	IVAVNKMDDK	SVK-----	-FSEERYNEI	KKELSGFLKK	V-GY-----	[182]
066_EF1a_Dimastigella_trypaniformis	DFIKNMITGT	SQADAAILII	ASGQGEFEAG	ISKD-----	--GQTKHEAL	LAYTLGVKQM	IVFVSNKMDDK	SVN-----	-YSQDRYNEI	VKEVGSYLYKK	V-GY-----	[168]

	120	130	140	150	160	170	180	190	200	210	220	
067_EF1a_Perkinsiella_amoebae	DFIKNMITGT	SQADAAILVI	ASGTGFEAG	ISGD-----	--GQTREHAL	LAYTLGVKQM	IVAVNKMDDK	TVN-----	-YQGRFDEI	KKEVQLYLKK	V-GY-----	[158]
068_EF1a_Leishmania_major	DFIKNMITGT	SQADAAILVI	DSTHGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVT-----	-YAQSRYEI	SKEVGAYLKR	V-GY-----	[183]
069_EF1a_Leishmania_infantum	DFIKNMITGT	SQADAAILMI	DSTHGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVT-----	-YAQSRYEI	SKEVGAYLKR	V-GY-----	[183]
070_EF1a_Leishmania_braziliensis	DFIKNMITGT	SQADAAILMI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVQ-----	-YSQARYEEI	SKEVGTYLKR	V-GY-----	[183]
071_EF1a_Leishmania_tarentolae	DFIKNMITGT	SQADAAILMI	DSTHGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVM-----	-YAQSRYEI	SKEVGAYLKR	V-GY-----	[183]
072_EF1a "L. tarentolae" FJ807257	DFIKNMITGT	SQADAAILMV	DSTGGFEAG	ISKN-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVS-----	-YSQARYEEI	SKEVGTYLKR	V-GY-----	[158]
073_EF1a_Leptomonas_costaricensis	DFIKNMITGT	SQADAAILMI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVQ-----	-YSQARYEEI	SKEVGAYLKR	V-GY-----	[99]
074_EF1a_Leptomonas_bifurcata	DFIKNMITGT	SQADAAILMI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVQ-----	-YSQARYEEI	SKEVGAYLKR	V-GY-----	[103]
075_EF1a_Leptomonas_podlipaevi	DFIKNMITGT	SQADAAILMI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVQ-----	-YAQARYEEI	TKEVGAYLKR	V-GY-----	[99]
076_EF1a_Herpetomonas_muscarum	DFIKNMITGT	SQADAAILMI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVCCNMDDK	TVG-----	-YSQARYEEI	MKEVSAYVKK	V-GY-----	[154]
077_EF1a_Herpetomonas_pestoi	DFIKNMITGT	SQADAAILMI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVQ-----	-FSEDYNEI	KKEVSTYLKK	V-GY-----	[110]
078_EF1a_Blastocritidia_culicis	DFIKNMITGT	SQADAAILMI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVCCNMDDK	TVN-----	-WSQRYNEI	TKEVGAYLKR	V-GY-----	[154]
079_EF1a_Trypanosoma_brucei	DFIKNMITGT	SQADAAILII	ASAQGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVN-----	-YGQERYDEI	KKEVSAYLKK	V-GY-----	[183]
080_EF1a_T. brucei equiperdum	DFIKNMITGT	SQADAAILII	ASAQGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVN-----	-YGQERYDEI	KKEVSAYLKK	V-GY-----	[157]
081_EF1a_Trypanosoma_aviu	DFIKNMITGT	SQADAAILII	ASAQGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	TVN-----	-YGQERYDEI	KKEVSAYLKK	V-GY-----	[165]
082_EF1a_Trypanosoma_cruzi	DFIKNMITGT	SQADAAILVI	ASAQGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	SVN-----	-FAQRYDEI	KKEVSAYLKK	V-GY-----	[183]
083_EF1a_Trypanosoma_rangeli	DFIKNMITGT	SQADAAILVI	ASAQGEFAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	VVCCNMDDK	SVN-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[183]
084_EF1a_Ploetia_costatum	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
085_EF1a_Ploetia_oblonga_cDNA	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
086_EF1a_Ploetia_oblonga_gDNA	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVACNMDDK	TVQ-----	-YNKDRYEEI	KKEVSAYLKK	V-GY-----	[182]
087_EF1a_Entosiphon_sp.	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVACNMDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
088_EF1a_Entosiphon_sulcatum	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVACNMDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[158]
089_EF1a_Peranema_tricophorum	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVAVNKMDDK	TVK-----	-YNKDRYEEI	KKEVSAYLKK	V-GY-----	[182]
090_EF1a_P. tricophorum_FJ807243	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVAVNKMDDK	TVK-----	-YNKDRYEEI	KKEVSAYLKK	V-GY-----	[180]
091_EF1a_Colacium_elongatum	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
092_EF1a_Colacium_mucronatum	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[164]
093_EF1a_Colacium_sp.	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
094_EF1a_Colacium_vesiculosum	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[168]
095_EF1a_Cryptoglena_pigra	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
096_EF1a_Cryptoglena_sukjai	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
097_EF1a_Discoplastis_adunca	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	SVN-----	-FAQRYDEI	KKEVSAYLKK	V-GY-----	[182]
098_EF1a_Euglena_clara	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[142]
099_EF1a_Euglena_deses	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
100_EF1a_Euglena_exilis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[142]
101_EF1a_Euglena_geniculata	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
102_EF1a_Euglena_gracilis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
103_EF1a_E. gracilis_X16890	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[183]
104_EF1a_E. gracilis_FJ807238	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[183]
105_EF1a_Euglena_granulata	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
106_EF1a_Euglena_hiemalis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
107_EF1a_Euglena_longa									-TQSYEEI	KKEVSAYLKK	V-GN-----	[21]
108_EF1a_Euglena_stellata	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[142]
109_EF1a_Euglena_viridis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	SVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[168]
110_EF1a_Eutreptia_pertyi	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
111_EF1a_Eutreptiella_pomquetensis					--QTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[52]
112_EF1a_Lepocinlis_acus_var._major	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
113_EF1a_Lepocinlis_ovum	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[142]
114_EF1a_Lepocinlis_spyrogyroides_clone_38												[0]
115_EF1a_L. spyrogyroides_clones_1_&_4	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
116_EF1a_Lepocinlis_steinii	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YNEARYEEI	KKEVSAYLKK	V-GY-----	[142]
117_EF1a_Lepocinlis_tripteris	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
118_EF1a_Monomorphina_aenigmatica	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
119_EF1a_Monomorphina_pyrum	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
120_EF1a_Monomorphina_sp.	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
121_EF1a_Phacus_orbicularis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[168]
122_EF1a_Phacus_similis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
123_EF1a_Strombomonas_acuminata	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[183]
124_EF1a_Strombomonas_borsteniensis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[142]
125_EF1a_Strombomonas_napiformis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
126_EF1a_Strombomonas_ovalis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
127_EF1a_Strombomonas_triquetra	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
128_EF1a_Trachelomonas_conradi	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
129_EF1a_Trachelomonas_echinata	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
130_EF1a_Trachelomonas_lefevrei	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[182]
131_EF1a_Trachelomonas_plancitonica	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSQARYEEI	KKEVSAYLKK	V-GY-----	[142]
132_EF1a_Trachelomonas_similis	DFIKNMITGT	SQADAAILVI	DSTGGFEAG	ISKD-----	--GQTREHAL	LAYTLGVKQM	IVATNKFDDK	TVK-----	-YSEARYEEI	KKEVSAYLKK	V-GY-----	[142]

	230	240	250	260	270	280	290	300	310	320	330	
001_aEF1a_Desulfurococcus_mobilis	-----D	-----VSKI	PFIPISAWTG	ENLIERSP	--NMPW--YN	G-----	-----	PTLVEALDTL	-EVPKPKINK	PLRIPIQDVF	[243]	
002_aEF1a_Sulfolobus_solfataricus	-----N	-----TNKV	RFVVPVAPAG	DNITHRSE	--NMKW--YN	G-----	-----	PTLEEYLDQL	-ELPPKPVDK	PLRIPIQDVF	[241]	
003_aEF1a_Thermoplasma_acidophilum	-----N	-----KDI	SFVPISGYKG	DNVTKPSP	--NMPW--YK	G-----	-----	PTLLQALDAF	-KVPEKPKINK	PLRIPVEDVF	[238]	
004_aEF1a_Halobacterium_sp.	-----N	-----PDDA	KFIATSAFEG	DNVSDHSD	--NTPW--YD	G-----	-----	PTLLEALNGL	-PVPQPPDTA	DLRLPIQDVF	[232]	
005_aEF1a_Methanocaldococcus_jannaschii	-----N	-----PDQI	DFIPTASLKG	DNVVKRSE	--NMPW--YK	G-----	-----	PTLVEALDKF	-IPPKPTNL	PLRIPIQDVF	[237]	
006_EF1a_Acanthamoeba_culbertsoni	-----N	-----PKNI	PFVPISGFHG	DNMVDKTD	--KMPW--YK	G-----	-----	PTLLEALDDI	-KPPKRPMDK	PLRVPDQVF	[231]	
007_EF1a_Neoparamoeba_branchipila	-----N	-----PAKI	PFVPISGWNG	DNMLERST	--NMNW--YK	G-----	-----	PTLLEALDSI	-IPPKRPIDK	ALRVPDQVF	[226]	
008_EF1a_Planoprotostelium_aurantium	-----N	-----PEKV	PFVPISGWNG	DNMLEKPD	--NMPW--WK	G-----	-----	PTLLEALDSI	-TEPKRPSDK	PLRVPDQVF	[225]	
009_EF1a_Physarum_polycephalum	-----N	-----PEKI	AFVPISGWNG	DNMLEKSA	--NLPW--YK	G-----	-----	PTLLEALDQI	-TEPKRPTDK	PLRVPDQVF	[225]	
010_EF1a_Amoebidium_parasiticum	-----N	-----PKAV	PFVPISGFNG	DNMVEPTD	--NMPW--YK	GWEVERKEG-	-----NATG	KTLLLEALDAI	-LPPKRPTDK	ALRLPLQDVF	[242]	
011_EF1a_Capsaspora_owczaraki	-----D	-----PKTV	AFVPISGWNG	DNMLEASE	--NMPW--FK	GWTIERKEG-	-----NASG	KTLLLEALDAI	-SPPKRPTDK	PLRLPLQDVF	[252]	
012_EF1a_Monosiga_ovata	-----N	-----PDTV	AFVPISGWNG	DNMIEASE	--KLPW--YK	GWEITRKDG-	-----NAKG	KTLLLEALDAI	-IPPERPTSK	PLRLPLQDVF	[245]	
013_EF1a_Salpingoeca_amphoridium	-----N	-----PDTV	AFVPISGWNG	DNMIEASE	--KLPW--YK	GWEVQRKDD-	-----AGGNAGK	KTLLLEALDAI	-HPPQRPTGK	PLRLPLQDVF	[246]	
014_EF1a_Ceanorhabditis_elegans	-----N	-----PKAV	PFVPISGFNG	DNMLEVSS	--NMPW--FK	GWAVERKEG-	-----NASG	KTLLLEALDSI	-IPPQRPTDR	PLRLPLQDVF	[254]	
015_EF1a_Drosophila_melanogaster	-----N	-----PAAV	AFVPISGWNG	DNMLEPST	--NMPW--FK	GWKVERKEG-	-----NADG	KTLLDALDAI	-LPPARPTDK	ALRLPLQDVF	[254]	
016_EF1a_Ornithorhynchus_anatinus	-----N	-----PATV	PFVPISGWNG	DNMLEPSP	--NMPW--FK	GWKVERKEG-	-----NASG	VSLLEALDTI	-LPTTRPTDK	ALRLPLQDVF	[254]	
017_EF1a_Xenopus_laevis	-----N	-----PATV	PFVPISGWNG	DNMLEAST	--NMPW--FK	GWKIERKEG-	-----NASG	ITLLEALDCI	-IPPQRPTNK	PLRLPLQDVF	[254]	
018_EF1a_Saccharomyces_cerevisiae	-----N	-----PKTV	PFVPISGWNG	DNMIEATT	--NAPW--YK	GWEKETKAG-	-----VVKG	KTLLLEAIDAI	-EQPSRPTDK	PLRLPLQDVF	[252]	
019_EF1a_Neuropora_crassa	-----N	-----PAGV	AFVPISGFNG	DNMLEPST	--NCPW--YK	GWEKETKAG-	-----KATG	KTLLLEAIDAI	-EPPKRPTDK	PLRLPLQDVF	[253]	
020_EF1a_Cryptococcus_neoformans	-----N	-----PKAV	PFVPISGWNG	DNMLEEST	--NMPW--YK	GWTKETKSG-	-----VSKG	KTLLLEAIDAI	-EPPTRPTDK	PLRLPLQDVF	[252]	
021_EF1a_Schizophyllum commune	-----N	-----PKTV	AFVPISGWNG	DNMLEEST	--NMPW--YK	GWTKETKAG-	-----VVKG	KTLLDAIDAI	-EPPVRPSDK	PLRLPLQDVF	[252]	
022_EF1a_Monoblepharis_insignis	-----N	-----PKTV	AFVPISGWNG	DNMIEATT	--NMPW--YK	GWEKETKSG-	-----KATG	KTLLVDAIDAI	-EPPTRPTDK	PLRLPLQDVF	[214]	
023_EF1a_Rhizophlyctis_harderi	-----N	-----PKAV	PFVPISGWNG	DNMLEPSE	--NTPW--FK	GWTKETKAG-	-----ALKG	TLLQAIDAI	-DLTPRPTDK	ALRLPLQDVF	[214]	
024_EF1a_Harpochytrium_sp.	-----N	-----PAAV	PFVPISGWNG	DNMIEASD	--NMPW--FK	SWQKETKGG-	-----KFTG	KTLLLEAIDSI	-EPPSRPTDK	PLRLPLQDVF	[198]	
025_EF1a_Absidia_glauca	-----N	-----PKSV	PFVPISGWNG	DNMLEEST	--NMPW--YK	GWNKETKAG-	-----AKSG	KTLLDAIDAI	-DPPQRPDSK	PLRLPLQDVF	[252]	
026_EF1a_Smittium_simulii	-----N	-----PKAV	AFVPISGFNG	DNMIEAST	--NMPW--YK	GWTKETKSG-	-----VSKG	VTLLDAIDAV	-EPPVRPSDK	PLRLPLQDVF	[243]	
027_EF1a_Chara_australis	-----N	-----PKDI	NFVPISGFEG	DNMIERST	--NMGW--YK	G-----	-----	PILLDALDLI	-SEPKRPSDK	PLRLPLQDVF	[225]	
028_EF1a_Physcomitrella_patens	-----N	-----PEKV	PFVPISGFEG	DNMIERST	--NLDW--YK	G-----	-----	PTLLEALDNV	-SEPKRPSDK	PLRLPLQDVF	[242]	
029_EF1a_Acetabularia_acetabulum	-----N	-----PEKI	NFVPISGFQG	DNMLEKSS	--NMNW--YK	G-----	-----	PTLLEALDMV	-EPPKRPSDK	PLRLPLQDVF	[225]	
030_EF1a_Chaetomorpha_coliformis	-----K	-----PDNV	PFIPISGWNG	DNMIERSS	--NMGW--YK	G-----	-----	PTLLEALDNV	-DPPKRPSDK	PLRLPLQDVF	[175]	
031_EF1a_Arabidopsis_thaliana	-----N	-----PKDI	PFVPISGFEG	DNMIERST	--NLDW--YK	G-----	-----	PTLLEALDQI	-NEPKRPSDK	PLRLPLQDVF	[242]	
032_EF1a_Oryza_sativa	-----N	-----PKDI	PFVPISGFEG	DNMIERST	--NLDW--YK	G-----	-----	PTLLEALDQI	-NEPNRPSDK	PLRLPLQDVF	[242]	
033_EF1a_Cyanidioschyzon_merolae	-----N	-----PKDV	PKVPISGWTG	DNLFERVPSD	HPLAKW--YK	G-----	-----	PALLEALDAI	-EPPKRPTDK	PLRLPLQDVF	[246]	
034_EF1a_Porphyrha_jezoensis	-----N	-----PAKV	PKVPTSGWTG	ENLFERTDKT	HALGKW--YK	G-----	-----	PCLLEALDNC	-DPPKRPSDK	PLRLPLQDVF	[246]	
035_EF1a_Plasmodium_falciparum	-----Q	-----ADKV	DFIPIISGFEG	DNLEIKSD	--KTPW--YK	G-----	-----	RTLLEALDTM	-EPPKRPSDK	PLRLPLQDVF	[240]	
036_EF1a_Cryptosporidium_parvum	-----N	-----TEKI	PFVAISGFVG	DNMVERSD	--KMPW--YK	G-----	-----	KTLLVEALDTM	-EPPKRPTDK	PLRLPLQDVF	[238]	
037_EF1a_Tetrahymena_pyriformis	-----K	-----PDTI	PFIPISGFNG	DNMLERST	--NAPW--YK	G-----	-----	PILVEALDAL	-EPPKRPSDK	PLRLPLQDVF	[243]	
038_EF1a_Stentor_coeruleus	-----N	-----PKDI	PFIPISGWL	DNMVERSP	--NLPW--YN	G-----	-----	PTLLGALDMI	-IPPKRHSDK	PLRLPLQDVF	[227]	
039_EF1a_Phytophthora_parasitica	-----K	-----PAKI	PFVPISGWEG	DNMIDRST	--NMPW--YK	G-----	-----	PYLLEALDNL	-NPPKRPSDK	PLRLPLQDVF	[242]	
040_EF1a_Mallomonas_rasilis	-----K	-----PAKI	NFVPISGWNG	DNMIDRS?	--NMAW--YK	G-----	-----	PFLLEALDLV	-NPPKRPTDK	PLRLPLQDVF	[229]	
041_EF1a_Phaeodactylum_tricornutum	-----K	-----PMKI	PFVPISGWEG	DNMVEKST	--NMAW--YK	G-----	-----	PYLLEALDSV	-TPPKRPTDK	ALRLPLQDVF	[242]	
042_EF1a_Thalassiosira_pseudonana	-----K	-----IDKV	KFIPISAWSG	ENLTERTK	--NMAW--YN	G-----	-----	PTLLEALDNV	-HPPKRPTDK	PLRLPLQDVF	[242]	
043_EF1a_Malawimonas_jakobiformis	-----N	-----LEKV	QFVPISGWNG	DNMMEKSP	--NMPW--YK	G-----	-----	FTLLEALDNL	-EPPKRPFEG	PLRLPLQDVF	[219]	
044_EF1a_Malawimonas_californiana	-----N	-----PEKI	PFVPISGWL	DNMIERSS	--NMGW--YK	G-----	-----	PTLLEALDAI	-EPPKRPSDK	PLRLPLQDVF	[242]	
045_EF1a_Trimastix_pyriformis	-----N	-----LEKV	QFVPISGFHG	DNMLEPSA	--NMPW--WK	G-----	-----	PTLLGALNLL	-EVPKRPVDK	PLRLPLQDVF	[228]	
046_EF1a_Dinenympha_exilis	-----N	-----PKDI	LMIPISGFNG	DNMLDRST	--NMPW--YK	D-----	-----	PILFDALDNL	-EVPKRPSDK	PLRLPLQDVF	[222]	
047_EF1a_Monocercomonoides_sp.	-----N	-----PKDI	PVPIPSGFQG	DNMLERSA	--NMPW--YK	G-----	-----	DILFDALDNL	-EVPKRPTDK	PLRLPLQDVF	[242]	
048_EF1a_Streblospio_strix	-----N	-----PKDI	QMIPIPSGFCG	DNMLEHSP	--NMPW--YK	G-----	-----	NTLFDLSDTL	-EVPKRPSDK	PLRLPLQDVF	[225]	
049_EF1a_Histomonas_meleagridis	-----K	-----PEQY	KFVPISGFVG	DNMTEKSP	--NMPW--YT	G-----	-----	GTLLLETDLT	-QPPKRPFDK	PLRLPLQDVF	[242]	
050_EF1a_Trichomonas_vaginalis	-----K	-----PEMF	RFVPISGWAG	DNMTEKSP	--NMPW--YN	G-----	-----	PYLLEALDSL	-QPPKRPFDK	PLRLPLQDVF	[242]	
051_EF1a_Hexamita_inflata	-----K	-----HWEF	DFVPLSGWTG	DNIMEASP	--KTPW--YK	G-----	-----	KCLTECIDGL	-KAPKRPNDK	PLRLPLQDVF	[226]	
052_EF1a_Spironucleus_vortens	-----K	-----KAPEF	NYIPTSGWTG	DNIMEKDT	--EKTW--YD	G-----	-----	PCLTEAIDSL	-KPPKRPTDK	CLRLPIQDVF	[226]	
053_EF1a_Andalucia_incarcerata	-----N	-----PAKV	PFVPSGWWG	DNMLERSP	--NMPW--WK	G-----	-----	PTLLEALDGL	-EPPSRPSDK	PLRLPLQDVF	[225]	
054_EF1a_Seculamonas_ecuadoriensis	-----N	-----PEKV	NFVPISGWNG	DNMLERSK	--NTSW--YT	G-----	-----	PTLVEAIDSF	-EPPKRPSDK	PLRLPLQDVF	[242]	
055_EF1a_Reclinomonas_americana	-----N	-----PKDV	NFIPISGWNG	DNMLERSP	--NTAW--YK	G-----	-----	PILTEAIDQF	-EPPKRPSDK	PLRLPLQDVF	[242]	
056_EF1a_Histiona_aroideis	-----N	-----PKDV	NFIPISGWNG	DNMLEKSA	--NMPW--YK	G-----	-----	PTLVEALDQF	-EPPKRPNDR	PLRVPDQVF	[242]	
057_EF1a_Jakoba_libera	-----N	-----PEQV	NFVPISGWNG	DNMLERSK	--NTPW--YT	G-----	-----	PTLLEAIDTF	-SEPKRPHDK	LRCPDQDVF	[242]	
058_EF1a_Stachyamoeba_lipophora	-----K	-----QEI-	PFIPISGFNG	DNMLDRST	--NMPW--YK	G-----	-----	PTLCEQLDNL	-EAPKRPTDK	PLRLPLQDVF	[241]	
059_EF1a_Naegleria_gruberi	-----Q	-----TEEGKPNVC	QFVPISGWTG	DNMIEKSD	--KMGW--YK	G-----	-----	PCLLDALDNL	-IEPVRPTDK	PLRLPLQDVF	[225]	
060_EF1a_Naegleria_andersoni	-----N	-----VDAADKPNLV	QFVPISGWTG	DNMIEKTD	--KMPW--YK	G-----	-----	PCLLDALDNL	-VEPVRPTDK	PLRLPLQDVF	[231]	
061_EF1a_Acrasis_rosea	-----N	-----PKNV	PFVPISGWAG	DNMLEKST	--NMPW--YK	G-----	-----	PTLLEALDAL	-EPPKRPTDK	PLRLPLQDVF	[225]	
062_EF1a_Sawyeria_marylandensis	-----K	-----PESV	PFVPISGWTG	DNMIEKSD	--NLKW--YT	G-----	-----	PTLLEALDMI	-KPPKPTDK	PLRLPLQDVF	[242]	
063_EF1a_Bodo_saltans	-----N	-----IEKV	RFVPISGWEG	DNMIEKSS	--RMEW--YK	G-----	-----	PTLLESDDL	-EPPTRPSDK	PLRLPLQDVF	[222]	
064_EF1a_Rhynchobodo_sp.	-----N	-----VEKV	KFVPISGWNG	DNMLESST	--NMPW--YK	G-----	-----	PTLLQALDGL	-EPPARPSDK	PLRLPLQDVF	[220]	
065_EF1a_Rhynchomonas_nasuta	-----N	-----VEKI	RFVPISGWNG	DNMIEKSD	--NMPW--YK	G-----	-----	PTLLEALDML	-EAPVRPSDK	PLRLPLQDVF	[241]	
066_EF1a_Dimastigella_trypaniformis	-----N	-----VEKV	KFIPISGWNG	DNMLEKSA	--NMPW--YK	G-----	-----	PTLLEALDSL	-EAPTRPHDK	PLRLPLQDVF	[227]	

	230	240	250	260	270	280	290	300	310	320	330					
067_EF1a_Perkinsiella_amoebae	-----	-----	N	-----	IEKV	QFVPISGWQG	DNMIERST	--NMNW--	YK	G	-----	PTLLEALDQL	-DPPVRPVDK	PLRLPLQDQV	[217]	
068_EF1a_Leishmania_major	-----	-----	N	-----	PEKV	RFIPISGWQG	DNMIEKSD	--NMPW--	YK	G	-----	PTLLDALDML	-EPPVRPVDK	PLRLPLQDQV	[242]	
069_EF1a_Leishmania_infantum	-----	-----	N	-----	PEKV	RFIPISGWQG	DNMIERSD	--NMPW--	YK	G	-----	PTLLDALDML	-EPPVRPVDK	PLRLPLQDQV	[242]	
070_EF1a_Leishmania_braziliensis	-----	-----	N	-----	PEKV	RFIPISGWQG	DNMIDKSE	--SMAW--	YK	G	-----	PTLLDALDML	-EAPVRPVDK	PLRLPLQDQV	[242]	
071_EF1a_Leishmania_tarentolae	-----	-----	N	-----	PEKV	RFIPISGWQG	DNMIEKSD	--NMPW--	YK	G	-----	PTLLDALDML	-EAPVRPVDK	PLRLPLQDQV	[242]	
072_EF1a "L. tarentolae" FJ807257	-----	-----	N	-----	IDKV	QFVPISGWNG	DNMIEKSE	--KMPW--	YK	G	-----	PTLLEALDML	-DQPVRLVDK	PLRLPLQDQV	[217]	
073_EF1a_Leptomonas_costaricensis	-----	-----	N	-----	PEKV	RFIPISGWQG	DNMIDKSD	--NMSW--	YK	G	-----	PTLLDALDML	-EAPVRPVDK	PLRLPLQDQV	[158]	
074_EF1a_Leptomonas_bifurcata	-----	-----	N	-----	PEKV	RFIPISGWQG	DNMIEKSD	--NMAW--	YK	G	-----	ATQLEALDML	-EAPVRPVDK	PLRLPLQDQV	[162]	
075_EF1a_Leptomonas_podlipaevi	-----	-----	N	-----	LEKV	RFIPISGWQG	DNMIEKSD	--NMPW--	YK	G	-----	PTQLESIDL	-EPPVRPVDK	PLRLPLQDQV	[158]	
076_EF1a_Herpetomonas_muscarum	-----	-----	N	-----	PEKV	KFIPIISGFQG	DNMIDRST	--NMSW--	YK	G	-----	PTLLEALDQL	-DPPVRPIEK	PLRLPLQDQV	[213]	
077_EF1a_Herpetomonas_pessoai	-----	-----	N	-----	PEKV	KFIPIISGWQG	DNMIEKST	--NMGW--	YK	G	-----	PTLLEALDSL	-DPPVRPSEK	PLRLPLQDQV	[169]	
078_EF1a_Blastocrithidia_culicis	-----	-----	N	-----	VEKV	RFIPISGWL	DNMIEKSD	--NMSW--	YK	G	-----	MTLLESIDL	-EPPVRLVDK	PLRLPLQDQV	[213]	
079_EF1a_Trypanosoma_brucei	-----	-----	N	-----	VEKV	RFVPIISGWQG	DNMIEKSE	--KMPW--	YK	G	-----	PTLLEALDML	-EPPVRPSDK	PLRLPLQDQV	[242]	
080_EF1a T. brucei equiperdum	-----	-----	N	-----	VEKV	RFVPIISGWQG	DNMIEKSE	--KMPW--	YK	G	-----	PTLLEALDML	-EPPVRPSDK	PLRLPLQDQV	[216]	
081_EF1a_Trypanosoma_avium	-----	-----	N	-----	VEKV	RFIPISGWQG	DNMIEKSE	--KMPW--	YK	G	-----	PTLLEALDML	-EPPVRPNDK	PLRLPLQDQV	[224]	
082_EF1a_Trypanosoma_cruzi	-----	-----	N	-----	VEKV	RFIPISGWQG	DNMIDKSE	--NMPW--	YK	G	-----	PTLLEALDML	-EPPVRPSDK	PLRLPLQDQV	[242]	
083_EF1a_Trypanosoma_rangeli	-----	-----	N	-----	VEKV	RFVPIISGWQG	DNMIEKSD	--NMPW--	YK	G	-----	PTLLEALDML	-EPPVRPSDK	PLRLPLQDQV	[242]	
084_EF1a_Ploeotia_costatum	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	NTLIQALDAL	-EPPKRPNDK	PLRLPLQDQV	[241]	
085_EF1a_Ploeotia_oblonga_cDNA	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	NTLIQALDAL	-EPPKRPNDK	PLRLPLQDQV	[241]	
086_EF1a_Ploeotia_oblonga_gDNA	-----	-----	N	-----	PDKV	PFVPIISGWNG	DNMLEKSD	--NMGW--	YK	G	-----	PTLLEALDNL	-ETPKRPADK	PLRLPLQDQV	[241]	
087_EF1a_Entosiphon_sp.	-----	-----	N	-----	PDKV	NFIPIISGWNG	DNMIEKSD	--NMAW--	YK	G	-----	PCLIEALDGL	-DAPKRPNDK	PLRLPLQDQV	[241]	
088_EF1a_Entosiphon_sulcatum	-----	-----	N	-----	PDKV	NFIPIISGWNG	DNMIEKSD	--NMAW--	YK	G	-----	PCLIEALDGL	-EAPKRPNDK	PLRLPLQDQV	[217]	
089_EF1a_Peranema_tricophorum	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEATE	--NMPW--	YK	G	-----	STLIDALDQL	-EPPKRPNDK	PLRLPLQDQV	[241]	
090_EF1a P. tricophorum_FJ807243	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEATE	--NMPW--	YK	G	-----	STLIDALDQL	-EPPKRPNDK	PLRLPLQDQV	[239]	
091_EF1a_Colacium_elongatum	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEASE	--NMGW--	YK	G	-----	PTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
092_EF1a_Colacium_mucronatum	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIDPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[223]	
093_EF1a_Colacium_sp.	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIESE	--NMGW--	YK	G	-----	PTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
094_EF1a_Colacium_vesiculosum	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIDPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[227]	
095_EF1a_Cryptoglena_pigra	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSE	--NKGW--	YK	G	-----	LTLIGALDAL	-EAPKRPNDK	PLRLPLQDQV	[241]	
096_EF1a_Cryptoglena_skujai	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSE	--NMAW--	YK	G	-----	LTLIGALDAL	-EPPKRPNDK	PLRLPLQDQV	[241]	
097_EF1a_Discoplastis_adunca	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSD	--NMTW--	YK	G	-----	PTLIGALDML	-DPPKRPNDK	PLRLPLQDQV	[241]	
098_EF1a_Euglena_clara	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEASE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[201]	
099_EF1a_Euglena_deses	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSD	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
100_EF1a_Euglena_exilis	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEASE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[201]	
101_EF1a_Euglena_geniculata	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	NTLIQALDAL	-EPPKRPNDK	PLRLPLQDQV	[241]	
102_EF1a_Euglena_gracilis	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSD	--NMGW--	YE	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
103_EF1a E. gracilis_X16890	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEASE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[242]	
104_EF1a E. gracilis_FJ807238	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSD	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[242]	
105_EF1a_Euglena_granulata	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSD	--NMSW--	YK	G	-----	NTLIQALDSL	-EPPKRPNDK	PLRLPLQDQV	[241]	
106_EF1a_Euglena_hiemalis	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEASE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
107_EF1a_Euglena_longa	-----	-----	N	-----	LEKV	RFIPISGWNG	DNMIEPSD	--NMGW--	YK	G	-----	LSLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[80]	
108_EF1a_Euglena_stellata	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EAPKRPNDK	PLRLPLQDQV	[201]	
109_EF1a_Euglena_viridis	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[227]	
110_EF1a_Eutreptia_pertyi	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSD	--NMSW--	YK	G	-----	NTLIQALDSL	-EPPKRPNDK	PLRLPLQDQV	[241]	
111_EF1a_Eutreptiella_pomquetensis	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	NTLIQALDAL	-EPPKRPNDK	PLRLPLQDQV	[111]	
112_EF1a_Lepocinlis_acus_var._major	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEASE	--NMPW--	YK	G	-----	MTLIGALDAL	-EPPKRPTDK	PLRLPLQDQV	[241]	
113_EF1a_Lepocinlis_ovum	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIDPSD	--NMPW--	YK	G	-----	MTLIGALDAL	-EPPKRPTDK	ALRLPLQDQV	[201]	
114_EF1a_Lepocinlis_spyrogyroides_clone_38	-----	-----	N	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	DK	ALRLPLQDQV	[12]
115_EF1a L. spyrogyroides_clones_1_&_4	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIDPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPTDK	PLRLPLQDQV	[241]	
116_EF1a_Lepocinlis_steinii	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSD	--NMKW--	YK	G	-----	PTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[201]	
117_EF1a_Lepocinlis_tripteris	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSD	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
118_EF1a_Monomorphina_aenigmatica	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	LTLIGALDAL	-EPPKRPNDK	PLRLPLQDQV	[241]	
119_EF1a_Monomorphina_pyrum	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIDPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPTDK	PLRLPLQDQV	[241]	
120_EF1a_Monomorphina_sp.	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPP?RPTDK	PLRLPLQDQV	[241]	
121_EF1a_Phacus_orbicularis	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPFD	--NMPW--	YK	G	-----	NTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[227]	
122_EF1a_Phacus_similis	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
123_EF1a_Strombomonas_acuminata	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEASE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[242]	
124_EF1a_Strombomonas_borysteniensis	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIDLSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[201]	
125_EF1a_Strombomonas_napiformis	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEASE	--NMTW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
126_EF1a_Strombomonas_ovalis	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIDASE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
127_EF1a_Strombomonas_triquetra	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEASE	--NMPW--	YK	G	-----	NTLIGALDAL	-EPPKRPNDK	PLRLPLQDQV	[241]	
128_EF1a_Trachelomonas_conradi	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEASE	--NMGW--	YK	G	-----	LTLIGALDNL	-KPPKRPNDK	PLRLPLQDQV	[241]	
129_EF1a_Trachelomonas_echinata	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEASE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
130_EF1a_Trachelomonas_lefevrei	-----	-----	Y	-----	PDKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[241]	
131_EF1a_Trachelomonas_planctonica	-----	-----	N	-----	PDKV	PFIPISGWNG	DNMIEPSE	--NMGW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[201]	
132_EF1a_Trachelomonas_similis	-----	-----	N	-----	PEKV	PFIPISGWNG	DNMIEASD	--NMKW--	YK	G	-----	LTLIGALDNL	-EPPKRPNDK	PLRLPLQDQV	[201]	

	230	240	250	260	270	280	290	300	310	320	330	
133_EFL1a_Trachelomonas_volvocina	-----	-----N	-----PDKV	PFIPISGWNG	DNMIEASE--	--NMPW--YK	G-----	-----	HTLIGALDNL	-EPPKRPSDK	PLRLPLQDVY	[241]
134_EFLenvironmental_sample_EU810232	-----	-----	-----AKSV	PVIPISGYMG	DNLIKPTST--	--NMDW--WK	GVDVKVGDK-	-----TVHV	HTLLDAFEKM	VTTPERSDKG	ALRVPISGVY	[266]
135_EFLCreolimax_fragrantissima	-----	-----	-----AKSV	PIIPISGWCG	DNLITKSD--	--KMAW--WS	GVDVLVGKE-	-----KIHV	HTLLDALEFM	VTIPTRTPDL	AMRMPLSGAF	[242]
136_EFLSpaeroforma_arctica	-----	-----	-----AKSV	PIIMPISWCG	DNLITKSN--	--NMDW--WK	GADVLVGKE-	-----KVHV	DCLLDCLCFM	VTIPVTRPDL	TMRMPLSGAF	[266]
137_EFLMonosiga_brevicollis	-----	-----	-----EESV	PVLPISGWKG	DNLITKTT--	--MSW--WN	GVDVVKVDE-	-----TIHI	DCLQDALEKM	VRVPQRATDK	PMRPISGVF	[266]
138_EFLXenopus_laevis	-----	-----	-----LKSV	PIIPISGWIG	DNLLTKSA--	--KMPW--WE	GVDVLVGKE-	-----KIHV	DTLFDALCEM	VQLPARPVDA	KLRLPLSGAF	[266]
139_EFLSpizellomyces_punctatus	-----	-----	-----QGSV	PVIPISGWIG	DNLLKKSE--	--KMSW--YQ	GQEVTLALSG-	-----KKVKV	HTLLDALNDF	AEMPERKND	PMRVPISGIY	[268]
140_EFLGaertneriomyces_semiglobiferous	-----	-----	-----QGCV	PVIPISGWQG	DNLLKKST--	--NMPW--YS	GQEVTLALSG-	-----KKIQV	HTLLDALNDF	AEMPERKND	PMRVPISGI-	[168]
141_EFLAllomyces_macrogynus	-----	-----	-----EGSV	PVIPISGWGM	DNLLKKSE--	--KMTW--WN	GQVKVLSLG-	-----KDIMV	VTLLNALNDF	AELPPRKSTA	AMRVPISGIY	[268]
142_EFLBlastocladiella_emersonii	-----	-----	-----EESV	PVIPISGWGM	DNLLKQST--	--NMTW--WA	GQEVKHPNG-	-----KKVKI	VTLLNALNDF	AELPARKSGS	PLRVPISGIY	[268]
143_EFLCatenophylctis_sp.	-----	-----	-----EQSV	PVIPISGWGM	DNLLKQST--	--NMTW--WK	GQEVVHPNG-	-----SKHTI	VTLLNALNDF	ADLPTRKSDA	ALRVPISGI-	[168]
144_EFLCoelomomyces_stegomyiae	-----	-----	-----EKCV	PILPISGWIG	DNLKEKSK--	--NMDW--WK	GQKVCNLLK-	-----KEVQV	DTLVDALNTF	AELPPRKSDA	AMRVPISGI-	[168]
145_EFLEntomophthoramuscae	-----	-----	-----EQNV	PVLPISGWGM	DNLLKKSE--	--KMTW--WK	GVDVKNQSG-	-----KDFHI	ETLQVALNDF	VELPSRVDDA	PMRLPVSGI-	[167]
146_EFLBasidiobolus_ranarum	-----	-----	-----EQSV	PVIPISGWQG	DNLLTKSA--	--KMPW--WK	GMDVVNQVG-	-----DKIHI	DTLL?ALNEM	VVLPTRNTEA	AMRLPVSGI-	[167]
147_EFLConidiobolus_coronatus	-----	-----	-----KDSV	PILPISGWQG	DNLLNQST--	--NMTW--WK	GVDITRIDG-	-----KKIHV	HTLKDALNDM	VTIPQRNVDA	PMRLPVSGI-	[167]
148_EFLTetraselmis_tetratele	-----	-----	-----LGCV	PILPISGWGM	DNLLSKSP--	--NMAW--WK	GVDISVGSE-	-----KLHI	ECLKDALNDM	SRLPDRNTTA	KLRLPVSGIY	[249]
149_EFLMesostigma_viride	-----	-----	-----AESV	PIIPISGWVG	DNLLKPSP--	--KMEW--WT	GMDVKNLGEG-	-----KTIKC	HTLLNVLNDF	ATVPERKTD	PMRLPVSGIY	[268]
150_EFLOstreococcus_tauri	-----	-----	-----EKS	PVLPISGWNG	DNLLKKSE--	--KMAW--WN	GVDVEGYDG-	-----KGAKFHI	DTLLNVLNDF	VDRPDNRTEA	QMLRPISGVY	[269]
151_EFLNephroselmis_olivacea	-----	-----	-----MDSV	PVMPISGWL	DNLLKQST--	--NMAW--WK	GADIDINKK-	-----KMKV	TTLLDCLNDM	VELPARDTAS	PMRLPVNGIY	[258]
152_EFLMicromonas_pusilla	-----	-----	-----EKCV	PVIPISGWQG	DNLIAKTS--	--NMGW--WS	GVDVIDYDL-	-----KGTKHHI	DTLLDVLNGM	VTIPKARNDA	DMRLPVSGVY	[252]
153_EFLParachlorella_kessleri	-----	-----	-----EKS	PIIPISGWGM	DNLITKST--	--MSW--WT	GVDVNNMLN-	-----KKIHV	HTLLDALNDF	AEPERKTD	KLRLPVSGIY	[253]
154_EFLHelicosporidium_sp.	-----	-----	-----EKCV	PVIPISGWGM	DNLIEKST--	--NMKW--WE	GVTVDVMDKD	GKKVEEKVKV	HTLLDALNDM	VRLPQRKIDA	PVRVPISGIY	[273]
155_EFLGonium_pectorale	-----	-----	-----KDSV	PVLPISGWL	DNLITKST--	--NMPW--YT	GQTVNLLKN-	-----EKIQV	HTLLDALNDF	VTIPERKIDA	KLRLPVSGIY	[267]
156_EFLChlamydomonas_reinhardtii	-----	-----	-----NKSV	PILPISGWL	DNLITKST--	--NMTW--YS	GQEVNLLKG-	-----EKIQV	HTLLDALNSF	VVPERKTD	KLRLPVSGIY	[267]
157_EFLScenedesmus_obliquus	-----	-----	-----EKS	PVLPISGWGM	DNLIKQST--	--NMAW--WK	GQEVVSADN-	-----RKQVI	NCLLDALNDM	VIVPERKTD	AMRVPISGAY	[267]
158_EFL"Oryza_sativa"	-----	-----	-----EKS	PIIPISGWGM	DNLIKSE--	--NMKW--WN	GMDVTGLNK-	-----EKVHI	ETLLNALNDM	VIVPERKTD	PMRVPISGAY	[267]
159_EFLChlorococcum_sp.	-----	-----	-----EASV	PFMPISGWGM	DNLIAKSE--	--KMGW--WS	GMDVKLSLG-	-----KEAQV	NTLLDYLNDV	VEVPERKTD	AMRVPISGIY	[250]
160_EFLOchlochaete_hystrix	-----	-----	-----EKCV	PILPISGWGM	DNLIAKST--	--NMTW--WN	GQEVNLLGEG-	-----KKIKV	TTVLDALNDF	AMPERKIDA	PMRVPISGIY	[253]
161_EFLUlvula_intestinalis	-----	-----	-----EGSV	PFLPISGWGH	DNLIAKSK--	--NMTW--WN	GQVKVSLAG-	-----KEVEV	VTLLNLYNEV	VELPARDTOK	KLRLPVSGIY	[250]
162_EFLUrospora_sp.	-----	-----	-----DNSV	PFLPISGWIG	DNLIKASA--	--NMTW--WK	PQDILTKLSG-	-----KPVQV	TTLLDYLNLV	VELPERDTK	KLRLPVSGIY	[201]
163_EFLGymnochlorella_stellata_ppc	-----	-----	-----DKSI	PFLPLSGFEG	DNLIQSK--	--NMDW--WN	GCDITTSWG-	-----DSIHV	NTLLDCLNSM	VRPPPRKIDA	KLRLPVAGVY	[279]
164_EFLLotharella_amoebiformis_ppc	-----	-----	-----EKS	PFLPLSGFEG	ENLLKSE--	--KMPW--WK	GNDVETSWG-	-----EKIHV	DTLLDCLNDM	VRPPPRKIDA	KLRLPVAGVY	[194]
165_EFLChlorarachnion_reptans_ppc	-----	-----	-----EKS	PFLPLSGFEG	DNLIEKSK--	--KMDW--WK	GNDIETVYG-	-----KKLHV	DTLYDCLDEM	AQPPERKIDA	KLRLPVAGVY	[203]
166_EFLChondrus_crispus	-----	-----	-----AKSV	PVIPISGWGM	DNLIT7SA--	--KMGW--WK	GCDVLVGKE-	-----KAHI	DTLSDALEKM	VRIPDRPTNA	KLRLPVSGVY	[266]
167_EFLGracilaria_changii	-----	-----	-----AKSV	PVVPISGWEG	DNLITKST--	--NMGW--WK	GCDVLVGKE-	-----KVHI	DTLSDALEKM	VKIPPRPTDK	KLRLPVSGVY	[266]
168_EFLCryptomonas_ovata	-----	-----	-----MDSV	PVVPISGWGM	DNLLKKSE--	--NMAW--WS	GVDVKAGNG-	-----TSVKI	VTLLNALNDF	AQPPVIRIDA	PMRCPISGIY	[252]
169_EFLGuillardia_theta	-----	-----	-----AASV	PVIPISGWGM	DNLLTKST--	--NMGW--WS	GVEVVPDGS-	-----TDKMKI	ETLLHALNDF	ARPPKRNDA	PMRCPISGIY	[267]
170_EFLRhodomonas_salina	-----	-----	-----AKSV	PVIPISGWGM	DNLLTKST--	--NMTW--WN	GIEIDVPDK-	-----GKMTI	VCLLDALNDM	AQPPQIRIDA	PMRCPISGIY	[266]
171_EFLPavlova_lutheri	-----	-----	-----ESGV	PVLPISGWGM	DNLLTKSA--	--KMDW--YT	GTDDVLIAGG-	-----KVHV	HTLYDALNDM	VKPERKTD	PMRLPVSGCY	[266]
172_EFLChrysochromulina.sp.	-----	-----	-----EKST	PVLPLSGWGM	DNLLKTE--	--MSW--WS	GQDILVEKA-	-----TMHI	DTLYDCLNDM	CKPPGRNDAA	AMRMPVSGIY	[206]
173_EFLIsochrysis_galbana	-----	-----	-----EQST	PVLPISGWGM	DNLLKKST--	--NMDW--YT	GQDILLPNK-	-----EKLHV	HTLYDCLNDM	CTPPGRNETA	PMRMPVSGIY	[267]
174_EFLEmiliana_huxleyi	-----	-----	-----AQSV	PVLPISGWGM	DNLEKTK--	--NMDW--WT	GCDVIGGPD-	-----GKTKMHV	DTLKECLNDF	CSPPVRTDKA	EMRMPVSGIY	[269]
175_EFLAchnanthes_kuwaitensis	-----	-----	-----PKKV	PFIPYSGFNG	DNLVEKSD--	--KASW--YK	GWTANINPK-	-----TKITG	FTLVECLNDF	ITPPKRFPDL	PVRLPISNIY	[240]
176_EFLAsterionella_glacialis	-----	-----	-----PKKV	PFIPYSGFNG	DNLVTKSD--	--KAPW--YK	GWTANTAPK-	-----VKVGT	FTLLDALNDF	ITPPKREPDA	KLRLPISVY	[240]
177_EFLDetonula_confervacea	-----	-----	-----PKKV	PFIPYSGFNG	DNLVKSD--	--KAPW--YK	GWTANTITPK-	-----KQVTG	FTIVDALNDF	IQPPIRHPEL	PVRLPISNIY	[240]
178_EFLThalassiosira_pseudonana	-----	-----	-----PKKV	PFIPYSGFNG	DNLVEKTD--	--KAPW--YN	GWEANINPK-	-----TVVKG	FTLVECLNDF	IQPPKRHPEL	PVRLPISNIY	[261]
179_EFLPerkinsus_marinus	-----	-----	-----EKCV	PVIPISGWGM	DNLLKKSE--	--KMSW--WK	GVDVTRDPA-	-----SKESFHI	DTLLDGLKDF	ATVPKRVTD	PMRMPVSGIY	[272]
180_EFLOxyrrhis_marina	-----	-----	-----DTNV	PFMPISGWCG	DNLFKEST--	--NMPW--YK	GQDVVPGT-	-----SEKIHV	HTLYDALDKF	CRVPVRPSEA	AMRMPVSGIY	[270]
181_EFLKarlodinium_micrum	-----	-----	-----EKNT	PYMPISGWGM	DNLLKKSE--	--NMDW--WK	GFDVAVGKE-	-----TIHV	DTMYDVLDM	CRVPERPMSA	PMRMPISGIY	[268]
182_EFLAmphidinium_carterae	-----	-----	-----KKN	PMLPISGWGM	DNLLKKST--	--NMAW--WS	GVDVEVGKE-	-----TLHI	DTLYEVLDM	CKIPERPVSA	PMRMPISGIY	[268]
183_EFLHeterocapsa_triquetra	-----	-----	-----EKNT	PVLPISGWGM	DNLLKKSD--	--NMAW--WK	GQDIEYGKE-	-----TIHV	DTAYDVLDM	CRVPERPMSA	PMRMPISGIY	[268]
184_EFLPlanoglabratella_operculis	-----	-----	-----TKKI	PFIPLSGFKG	DNLAEPSK--	--NMAW--YK	GFEIKMKDK-	-----NKNKIVVKG	HTLIDALEKV	VKPPKRANKK	AFRMPVSGVY	[214]
185_EFLReticulomyxa_filosa	-----	-----	-----TKKI	PFIPLSGFKG	ENLTKVSE--	--NMPW--WK	GFSVQIEKD-	-----TVTGT	LTMDALEKT	VNQPKRPTDK	PMRMPVSGVY	[238]
186_EFLLotharella_amoebiformis	EK-----	-----GKKKG-	-----PMPI	PVIPISGWGT	DNLIQST--	--KMPW--FK	RWKATAPNK-	-----KEYTG	KTLFECLDQF	VQPVARKLKD	KLRLPVSGVY	[290]
187_EFLGymnochlorella_stellata	KKDDKGGKKK	GKGEKKKAG-	-----PNLI	PVIPISGWCG	DNLIIPST--	--KMPWFNKK	GWTATTPSG-	-----VKVKG	QTLFEALDQF	VEPVTRDLSK	KLRLPVSGVY	[276]
188_EFLChlorarachnion_reptans	-----	-----KKKKG-	-----PNLV	PVIPISGWGM	DNLIKPTST--	--NMPWFNKK	GWTATTPSG-	-----VKVKG	QTLFEALDQF	VEPVTRDLSK	KLRLPVSGVY	[253]
189_EFLDiplonema_ambulator	-----	-----	-----ENSV	PVLPISGWIG	DNLLTKSA--	--KMPW--WK	GQDVIVGAD-	-----EKLHV	ETLVQALDM	VRMPERKND	PMRMPVSGVY	[134]
190_EFLDiplonema_papillatum	-----	-----	-----EQCV	PVIPISGWIG	DNLLKKSE--	--KMPW--WK	GQDVIVGAD-	-----EKLHV	ETLVQALDM	VRMPERKND	PMRMPVSGVY	[258]
191_EFLRhynchopus_euleeides	-----	-----	-----EKS	PVLPISGWGM	DNLLKKSD--	--KMAW--WK	GQDVIVGAD-	-----KSNLHI	ETLVQALDM	VRMPERKND	PMRMPVSGVY	[139]
192_EFLPetalomonas_cantuscygni	-----	-----	-----EKS	PVIPISGWGM	DNLLKKSE--	--KMAW--WN	GQDVIVGAD-	-----SKLHA	DTLLDCLDKY	VKLPPRQTDK	PMRMPVSGMY	[215]
193_EFLTrypanoplasma_borrel	-----	-----	-----ENEI	PIIPMAGFHG	ENMLKKSD--	--KMPW--WN	GVDVVKVAA-	-----GGTVVHV	DCLLDALNDM	VNLPDRNTA	PLRAPVSGCL	[275]
194_EFLNeobodo_salien	-----	-----	-----EKEI	PILPLAGFHG	ENMLKASD--	--KMPW--WK	GVTVFDPST-	-----KKDIKV	NCLLDALNDM	VNLPDRNTEG	ALRAPVSGVY	[256]

	340	350	360	370	380	390	400	410	420	430	440
001_aEF1a_Desulfurococcus_mobilis	-NISGIGVVP	VGRVETGVLM	VGDKLVFMPA	GLVA-----	---EVKTIET	HHTKIEKAEP	GDNIGFNVKG	VEK-KDIKRG	DVAGSLD---	V--P-PTVA-	---DEFTARI
002_aEF1a_Sulfolobus_solfataricus	-SISGVGTVP	VGRVESGVLM	VGDKIVFMPA	GKVG-----	---EVRSIET	HHTKMDKAEP	GDNIGFNVRG	VEK-KDIKRG	DVVGHPN---	N--P-PTVA-	---DEFTARI
003_aEF1a_Thermoplasma_acidophilum	-SITGIGTVP	VGRVETGVLM	PGDKVIFLPA	DKQG-----	---DVKSIEM	HHEPLQQAEP	GDNIGFNVRG	IAK-NDIKRG	DVCGHLD---	T--P-PTVV-	---KAFTAQI
004_aEF1a_Halobacterium_sp.	-TISGIGTVP	VGRITETGMN	TGDNVSFQPS	DVGG-----	---EVKTIEM	HHEEVPNAEP	GDNVGFNVRG	IGK-KDIRRG	DVCGPAD---	D--P-PSVA-	---DTFQAQV
005_aEF1a_Methanocaldococcus_jannaschii	-SITGVGTVP	VGRVETGVLM	PGDKVVFAPA	GVSG-----	---EVKSIEM	HHEQTPQAEP	GDNIGFNVRG	VSK-KDIKRG	DVCGHPD---	N--P-PTVA-	---EEFTAQI
006_EF1a_Acanthamoeba_culbertsoni	-KIGGIGTVP	VGRITETGILK	PGMVVTFAPV	NVTT-----	---EVKSSEM	HHEAMPEAVP	GDNVGFNVKN	VSI-KDIRRG	NVCGDSK---	K-DP-PQET-	---EDFTAQV
007_EF1a_Neoparamoeba_branchipila	-KIGGIGTVP	VGRVETGVLM	PGMVVTFAPP	MITT-----	---EVKSSEM	HHEQMKAQVP	GDNVGFNVKN	VSV-KDIRRG	NVAGDSK---	N-DP-PAGA-	---EDFTAQV
008_EF1a_Planoprotostelium_aurantium	-KIGGIGTVP	VGRVETGILK	PGSVVTFAPA	GVTT-----	---EVKSSEM	HHVQLPEARP	GDNVGFNVKN	VSV-KDIRRG	YVAGDAK---	A-DP-PLGT-	---EFFVAQV
009_EF1a_Physarum_polycephalum	-KIGGIGTVP	VGRVETGILK	PGMIVTFAPA	NLST-----	---EVKSSEM	HHVALPEAVP	GDNVGFNVKN	LSV-KDIRRG	MVAGDSK---	N-SP-PARA-	---ESFNAQV
010_EF1a_Amoebidium_parasiticum	-KIGGIGTVP	VGRVETGVIK	PGMVVTFAPT	GVTT-----	---EVKSSEM	HHEALSE---	---	---	---	---	---
011_EF1a_Capsaspora_owczaraki	-KIGGIGTVP	VGRVETGVLM	PGMVVTFAPS	NVTT-----	---EVKSSEM	HHESLPEANP	GDNVGFNVKN	VAV-KDIRRG	NVAGDSK---	N-DP-PKET-	---KTFTAQV
012_EF1a_Monosiga_ovata	-KIGGIGTVP	VGRVETGTLK	PGMVVTFAPG	NVTT-----	---EVKSSEM	HHEQLQEAVP	GDNVGFNIKN	VSV-KDIRRG	NVCGDSK---	N-DP-PKAS-	---ESFLAQV
013_EF1a_Salpingoeca_amphoridium	-KIGGIGTVP	VGRVETGVLM	PGMVVTFAPG	NVST-----	---EVKSSEM	HHEQLVEAVP	GDNVGFNVKN	LSV-KDIRRG	NVCGDSK---	N-DP-PKAT-	---EKFNAQV
014_EF1a_Ceanorhabditis_elegans	-KIGGIGTVP	VGRVETGIIK	PGMVVTFAPQ	NVTT-----	---EVKSSEM	HHESLPEAVP	GDNVGFNVKN	VSV-KDIRRG	SVCSDSK---	Q-DP-AKEA-	---RTFHAQV
015_EF1a_Drosophila_melanogaster	-KIGGIGTVP	VGRVETGVLM	PGTVVVFAPA	NITT-----	---EVKSSEM	HHEALQEAVP	GDNVGFNVKN	VSV-KELRRG	YVAGDSK---	A-NP-PKGA-	---ADFTAQV
016_EF1a_Ornithorhynchus_anatinus	-KIGGIGTVP	VGRVETGILR	PGMVVTFAPV	NITT-----	---EVKSSEM	HHEALSEALP	GDNVGFNVKN	VSV-KDIRRG	NVCGDSK---	S-DP-PQEA-	---AQFTSQV
017_EF1a_Xenopus_laevis	-KIGGIGTVP	VGRVETGVLM	PGMIVTFAPS	NVTT-----	---EVKSSEM	HHEALQEAAP	GDNVGFNVKN	ISV-KDIRRG	NVAGDSK---	N-DP-PMQA-	---GSFTAQV
018_EF1a_Saccharomyces_cerevisiae	-KIGGIGTVP	VGRVETGVIK	PGMVVTFAPA	GVTT-----	---EVKSSEM	HHEQLQGVVP	GDNVGFNVKN	VSV-KEIRRG	NVCGDAK---	N-DP-PKGC-	---ASFNAQV
019_EF1a_Neurospora_crassa	-KIGGIGTVP	VGRITETGVLM	PGMVVTFAPS	NVTT-----	---EVKSSEM	HHEQLAQGVVP	GDNVGFNVKN	VSV-KDIRRG	NVAGDSK---	N-DP-PAGA-	---ASFTAQV
020_EF1a_Cryptococcus_neoformans	-KIGGIGTVP	VGRVETGVIK	AGMVVVFAPT	NVTT-----	---EVKSSEM	HHEQTPGLP	GDNVGFNVKN	VSI-KDIRRG	NVCGDSK---	N-DP-PMEA-	---ASFNAQV
021_EF1a_Schizosaccharomyces_commune	-KIGGIGTVP	VGRVETGIIK	AGMVVVFAPS	NVTT-----	---EVKSSEM	HHEQLAEGKP	GDNVGFNVKN	VSV-KDIRRG	NVAGDSK---	N-DP-AKEA-	---ASFNAQV
022_EF1a_Monoblepharis_insignis	-KIGGIGTVP	VGRVETGVIK	PGMVVTFAPS	NVTT-----	---EVKSSEM	HHESLAEGLP	GDNVGFNVKN	VSV-KDIRRG	NVAGDSK---	N-DP-AQES-	---ASFNAQV
023_EF1a_Rhizophlyctis_harderi	-KIGGIGTVP	VGRVETGTIK	AGMIVNFSPA	NVTT-----	---EVKSSEM	HHELTLEAGL	GDNVGFNIKN	VSV-KEIRRG	MVCSDTK---	N-DP-AKEA-	---ASFNAQI
024_EF1a_Harpochytrium_sp.	-KIGGIGTVP	VGRVETGIIK	PGMVVTFAPT	GVTT-----	---EVKSSEM	HHEALPEALP	GDNVGFNVKN	VSV-KDIRRG	NVAGDSK---	N-DP-PRET-	---ASFNAQV
025_EF1a_Absidia_glauca	-KIGGIGTVP	VGRVETGVIK	AGMVVVFAPA	NVTT-----	---EVKSSEM	HHEQLVEGLP	GDNVGFNVKN	VSV-KDIRRG	NVCGDSK---	N-DP-AKEA-	---GSFTAQV
026_EF1a_Smittium_simulii	-KIGGIGTVP	VGRVETGVIK	AGMIVTFAPS	YVTT-----	---EVKSSEM	HHETLTEGLP	GDNVGFNIKN	VSV-KDIRRG	NVCGDSK---	N-DP-PKET-	---GTFTAQV
027_EF1a_Chara_australis	-KIGGIGTVP	VGRVETGIIK	PGMVVTFAPS	GLTT-----	---EVKSSEM	HHEAMTEALP	GDNVGFNVKN	VSV-KELKRG	FVAGDSK---	N-DP-AKEA-	---ASFTSQV
028_EF1a_Physcomitrella_patens	-KIGGIGTVP	VGRVETGIIR	PGMLVFCAPT	GLTT-----	---EVKSSEM	HHESMPEAHP	GDNVGFNVKN	VAV-KDLKRG	YVAGDSK---	N-DP-AKEA-	---ANFTAQV
029_EF1a_Acetabularia_acetabulum	-KIGGIGTVP	VGRVETGVIK	PGMIVTFAPT	GLTT-----	---EVKSSEM	HHEALPEAVP	GDNVGFNVKN	VSV-KEIKRG	YVAGDSK---	S-DP-AKGC-	---DNFTAQV
030_EF1a_Chaetomorpha_coliformis	-KIGGIGTVP	VGRVETGVIK	PGMVVVFAPV	GLTT-----	---EVKSSEM	HHEALPEAGP	GDNVGFNVKN	VSV-KELKRG	FVAGDSK---	N-DP-AKGC-	---QYFKAQV
031_EF1a_Arabidopsis_thaliana	-KIGGIGTVP	VGRVETGMIK	PGMVVTFAPT	GLTT-----	---EVKSSEM	HHEALLAALP	GDNVGFNVKN	VAV-KDLKRG	YVAGDSK---	D-DP-AKGA-	---ANFTSQV
032_EF1a_Oryza_sativa	-KIGGIGTVP	VGRVETGVLM	PGMVVTFGSP	GLTT-----	---EVKSSEM	HHEALQEAAP	GDNVGFNVKN	VAV-KDLKRG	YVAGDSK---	D-DP-AKEA-	---ASFTSQV
033_EF1a_Cyanidioschyzon_merolae	-KIGGIGTVP	VGRVETGILR	PGMVVTFAPT	GLTT-----	---EVKSSEM	HHESLPEAGP	GDNVGFNVKN	VSV-KELKRG	FVCGDSK---	N-DP-PKAA-	---EDFKAQV
034_EF1a_Porphyrha_jezoensis	-KIGGIGTVP	VGRVETGLIK	PGMVVTFAPS	GLST-----	---EVKSSEM	HHEALPQAGP	GDNVGFNVKN	VSV-KDLKRG	YVCGDSK---	N-DP-PKGC-	---ASFNAQV
035_EF1a_Plasmodium_falciparum	-KIGGIGTVP	VGRVETGILK	AGMVLNFAPS	AVVS-----	---ECKSSEM	HKEVLEEARP	GDNIGFNIKN	VSV-KEIKRG	YVAGDSK---	N-EP-AKGC-	---SKFTAQV
036_EF1a_Cryptosporidium_parvum	-KIGGIGTVP	VGRVETGIIR	PGMVVVFAPA	GVTT-----	---EVKSSEM	HHEQMPEAVP	GDNVGFNVKN	VSI-KDIKRG	FVAGDSK---	N-DP-AKGC-	---EDFTAQV
037_EF1a_Tetrahymena_pyriformis	-KIGGIGTVP	VGRVETGVIK	PGMSIQFAPN	KVIA-----	---ECKSSEM	HHEQLPEAVP	GDNVGFNIKN	VSV-KDIRRG	NVAGDSK---	N-DP-AKEA-	---ATFYSQV
038_EF1a_Stentor_coeruleus	-KISGIGTVP	VGRVETGSIK	AGMVVVFAPS	GVTT-----	---EVKSSEM	HHETLEVANP	GDNVGFNVKN	VSV-KDIKRG	YVAGDSK---	A-DP-AKEC-	---ENLTAQV
039_EF1a_Phytophthora_parasitica	-KIGGIGTVP	VGRVETGVIK	PGMVATFGPV	GLST-----	---EVKSSEM	HHEALPEAVP	GDNVGFNVKN	VSV-KELRRG	YVAGDSK---	N-DP-AKGT-	---QDFTAQV
040_EF1a_Mallomonas_rasilis	-KIGGIGTVP	VGRVETGILK	PNMVVTFAPV	GLTT-----	---EVKSSEM	HHEALAEALP	GDNVGFNVKN	VSV-KDIRRG	SVASNSA---	E-DP-AKGC-	---DTFTAQV
041_EF1a_Phaeodactylum_tricornutum	-KIGGIGTVP	VGRVETGVIK	PGIHAMFAPS	GIIA-----	---EIKSSEM	HHESLPEAVP	GDNVGFNVKN	VAV-KDLRRG	FVAGDSK---	A-SP-ASGV-	---SSFEAQV
042_EF1a_Thalassiosira_pseudonana	-KIGGIGTVP	AGRVETGVMM	PGMHIEFAPT	GIVA-----	---DVKSIEM	HHQIPEALP	GDNIGFNIKN	VAV-QDLHRG	DVASEAG---	K-NE-ATAA-	---KSFDAQV
043_EF1a_Malawimonas_jakobiformis	-KIGGIGTVP	VGRVETGILK	PGMVVQFAPS	GVTT-----	---EVKSSEM	HHEAMPEAVP	GDNVGFNVKN	VAV-KDLRRG	FVCGDTK---	N-DP-PMEA-	---GEFTAQV
044_EF1a_Malawimonas_californiana	-KIGGIGTVP	VGRVETGIIK	PGMVVTFAPV	GVTT-----	---EVKSSEM	HHESMPEAQP	GDNVGFNVKN	VSV-KDIRRG	NVCGDSK---	N-DP-PKES-	---ADFTAQV
045_EF1a_Trimastix_pyriformis	-KIGGIGTVP	VGRVETGILR	PGQVITIAPA	NLST-----	---EVKSSEM	HHTALEQATP	GDNVGFNVKN	IAV-KDLKRG	FVVGDSR---	V-DP-PQET-	---EKFVAQV
046_EF1a_Dinenympha_exilis	-KIGGIGTVP	VGRVETGILT	PGQVITIAPC	MITT-----	---EVKSSEM	HHEALTQAVP	GDNVGFNVKG	IAV-KDLKRG	FVAGDAK---	N-DP-PKEA-	---DTFSAQV
047_EF1a_Monocercomonoides_sp.	-KIGGIGTVP	VGRVETGVLT	PGMVVITIAPA	AITT-----	---EVKSSEM	HHEALRAVP	GDNVGFNVKN	ISV-KDIRRG	NVAGDSK---	Q-DP-PMEA-	---ESFVAQV
048_EF1a_Streblospio_strix	-KIGGIGTVP	VGRVETGVLT	PGQVITIAPA	MITT-----	---ECKSSEM	HHEALTQAVP	GDNVGFNLKG	VSV-KDIKRG	YVCGDSK---	Q-DP-PQET-	---ESSQAQV
049_EF1a_Histomonas_meleagridis	-KISGIGTVP	VGRVESGIMK	PNMTVVFAPS	TVTT-----	---EVKSSEM	HHTLQPEAVP	GDNIGFNVKN	VAV-SDIKRG	YVVGAN---	R-DP-PMQC-	---INFTAQV
050_EF1a_Trichomonas_vaginalis	-KINGIGTVP	VGRVESGTMK	PGMIVNFAPS	TVTA-----	---EVKSSEM	HHESLPEALP	GDNIGFNVKN	VST-ADVVRG	YVVGDTK---	R-DP-PVEC-	---ASFTAQV
051_EF1a_Hexamita_inflata	-KINGVGTVP	AGRVESGELI	PGMVVVFAPA	GKGT-----	---EVKSSEM	HHEQLAKAGP	GDNVGFNIKN	LSA-KDIKRG	YVVGDNV---	N-DA-PKGC-	---EYFKANV
052_EF1a_Spiroplasma_vortens	-KINGVGTVP	AGRVESGCLL	PNTLAVFAPT	NNTA-----	---EVKSSEM	HHEELPQAEV	GDNVGFNVKG	IAA-KDIKRG	YVVGDSK---	ESDP-PGRV-	---KSFEAQV
053_EF1a_Andalucia_incarcerata	-KIGGIGTVP	VGRVETGKLE	AGMTVVFAPS	GVTT-----	---EVKSSEM	HHEALEEALP	GDNVGFNVKN	VSV-KDIRRG	NVAGDSK---	N-DP-AKPC-	---ESFVAQV
054_EF1a_Seculamonas_ecuadoriensis	-KIGGIGTVP	VGRVETGIMK	PGMTVLFAPS	GLST-----	---EVKSSEM	HHTLQPEAVP	GDNVGFNIKN	IAV-KDIRRG	YVAGDSK---	N-DP-AKEA-	---ASFNAQV
055_EF1a_Reclinomonas_americana	-KIGGIGTVP	VGRVETGILR	PGMVVTFAPG	GLST-----	---EVKSSEM	HHTLQPEALP	GDNVGFNVKN	VSV-KDIRRG	YVAGDSK---	N-DP-AKEA-	---TSFTAQV
056_EF1a_Histiona_aroidea	-KIGGIGTVP	VGRVETGVMM	PGTTVVFAPC	GLST-----	---EVKSSEM	HHTLQPEAVP	GDNVGFNVKN	LSV-KDIRRG	YVAGDSK---	N-DP-AKEA-	---ESFTAQV
057_EF1a_Jakoba_libera	-KIGGIGTVP	VGRVETGIMK	PGMIVTFAPS	GVTT-----	---ECKSSEM	HHTLQPEALP	GDNVGFNVKG	VSV-KDIRRG	NVAGDSK---	N-DP-AKES-	---TSFNAQV
058_EF1a_Stachyamoeba_lipophora	-KISGIGTVP	VGRVETGILK	PGMVVTFAPG	GVST-----	---DVKSSEM	HHEQLPEALP	GDNVGFNVKN	ISV-KDVVRG	FVAGETK---	N-DP-PKES-	---ADFTAQV
059_EF1a_Naegleria_gruberi	-KIGGIGTVP	VGRVETGKLM	PGMVMHFAPG	QAADT-----	---EVKSSEM	HHTSVPAGP	GDNVGFNVKG	LST-TDIKRG	YVAGDSK---	N-DP-SREA-	---VSFNAQV
060_EF1a_Naegleria_andersoni	-KIGGIGTVP	VGRVETGKLM	PGMIMHFAPG	QAADT-----	---EVKSSEM	HHTSVPAGP	GDNVGFNVKG	LST-TDIKRG	YVAGDSK---	S-DP-SREA-	---VSFNAQV
061_EF1a_Acrasis_rosea	-KIGGIGTVP	VGRVETGVMM	PGDVLVFAV	DVST-----	---EVKSSEM	HHEQLAEAIL	GDNVGFNIKN	VAV-KDIRRG	NVCSNIAK---	N-DP-ARGC-	---ENFEAQV
062_EF1a_Sawyeria_marylandensis	-KIGGIGTVP	VGRVETGVMT	PGMTVVFSPA	AKST-----	---EVKSSEM	HHTSIPQAEV	GDNIGFNVKN	LSV-KDIKRG	NVCGDIK---	N-DP-PKEA-	---ASFVAQV
063_EF1a_Bodo_saltans	-KIGGIGTVP	VGRVETGILR	PGDVTTFAPA	NITT-----	---EVKSSEM	HHESLPEAIP	GDNVGFNVKN	LSI-KDIRRG	FVCGSAK---	V-DP-PKEC-	---ESFTAQV
064_EF1a_Rhynchobodo_sp.	-KIGGIGTVP	VGRVETGIMK	PGDVTTFAPA	GVTT-----	---EVKSSEM	HHESLPEATP	GDNVGFNVKN	ISV-KDVVRG	FVCGNTK---	Q-DP-PKGC-	---DFFTAQV
065_EF1a_Rhynchomonas_nasuta	-KIGGIGTVP	VGRVETGIMK	PGDVTTFAPA	NVST-----	---EVKSSEM	HHESLPEATP	GDNIGFNVKN	LSV-KDIRRG	FVCSNSK---	N-DP-AKEA-	---ENFLAQV
066_EF1a_Dimastigella_trypaniformis	-KIGGIGTVP	VGRVETGILK	PGDVTTFAPA	GVTT-----	---EVKSSEM	HHEQLAEAAP	GDNVGFNVKN	ISI-KDIRRG	FVCSNSK---	N-DP-AKEA-	---ENFLAHV

	340	350	360	370	380	390	400	410	420	430	440	
067_EF1a_Perkinsiella_amoebae	-KIGGIGTVP	VGRVETGVL	PGDVVTFAPT	NLTT-----	--EVKSVM	HHEALPEATP	GMNIGFNVKN	VSV-KDIRRG	FVCSNSK---	Q-DP-AKEC	--EHFIA--	[305]
068_EF1a_Leishmania_major	-KIGGIGTVP	VGRVETGIMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[332]
069_EF1a_Leishmania_infantum	-KIGGIGTVP	VGRVETGIMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[332]
070_EF1a_Leishmania_braziliensis	-KIGGIGTVP	VGRVETGIMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[332]
071_EF1a_Leishmania_tarentolae	-KIGGIGTVP	VGRVETGIMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[332]
072_EF1a "L. tarentolae" FJ807257	-KIGGIGTVP	VGRVETGVMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTA--	[305]
073_EF1a_Leptomonas_costaricensis	-KIGGIGTVP	VGRVETGIMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	-----	-----	-----	-----	-----	-----	[198]
074_EF1a_Leptomonas_bifurcata	-KIGGIGTVP	VGRVETGIMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[252]
075_EF1a_Leptomonas_podlipaevi	-KIGGIGTVP	VGRVETGIMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[248]
076_EF1a_Herpetomonas_muscarum	-KIGGIGTVP	VGRVETGVMK	PGDVVFFAPS	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCSNSK---	N-DP-AREA	--EDFTAQV	[303]
077_EF1a_Herpetomonas_pestoi	-KIGGIGTVP	VGRVETGVMR	PGDVVVFAPA	NVTT-----	--EVKSIEM	HHEQLAAEQP	GDNVGFNVKN	VSV-KDIRRG	NVCSNAK---	N-DP-AREA	--EDFTA--	[257]
078_EF1a_Blastocirithidia_culicis	-KIGGIGTVP	VGRVETGVMN	PGDVVTFAPA	NVTT-----	--EVKSIEM	HH-----	-----	-----	-----	-----	-----	[255]
079_EF1a_Trypanosoma_brucei	-KIGGIGTVP	VGRVETGVMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAEATP	GDNVGFNVKN	VSV-KDIRRG	NVCGNTK---	N-DP-PKEA	--ADFTAQV	[332]
080_EF1a T. brucei equiperdum	-KIGGIGTVP	VGRVETGVMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HH-----	-----	-----	-----	-----	-----	[258]
081_EF1a_Trypanosoma_avium	-KIGGIGTVP	VGRVETGVMR	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAEATP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[314]
082_EF1a_Trypanosoma_cruzi	-KIGGIGTVP	VGRVETGIMK	PGDVVTFAPA	NVTT-----	--EVKSIEM	HHEQLAEATP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[332]
083_EF1a_Trypanosoma_rangeli	-KIGGIGTVP	VGRVETGVMK	PGDVVTFSPA	NVTT-----	--EVKSIEM	HHEQLAEATP	GDNVGFNVKN	VSV-KDIRRG	NVCGNSK---	N-DP-PKEA	--ADFTAQV	[332]
084_EF1a_Ploeotia_costatum	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--ADFTAQV	[331]
085_EF1a_Ploeotia_oblonga_cDNA	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--ADFTAQV	[331]
086_EF1a_Ploeotia_oblonga_gDNA	-KIGGIGTVP	VGRVETGVMK	PGDVVTFAPN	NLTT-----	--EVKSIEM	HHEALPEAVP	GDNVGFNVKN	VAV-KDIRRG	FVCSNSK---	Q-DP-AKEA	--EDFTAQV	[331]
087_EF1a_Entosiphon_sp.	-KIGGIGTVP	VGRVETGILK	PGDVVTFAPV	NLTT-----	--EVKSIEM	HHEALPEAVP	GDNVGFNVKN	VSV-KEIRRG	FVASNAK---	N-DP-AKET-	--ETFTAQV	[331]
088_EF1a_Entosiphon_sulcatum	-KIGGIGTVP	VGRVETGILK	PGDVVTFAPV	NLTT-----	--EVKSIEM	HHEALPEAVP	GDNVGFNVKN	VSV-KEIRRG	FVASNAK---	N-DP-AKET-	--ETFTA--	[305]
089_EF1a_Peranema_tricophorum	-KIGGIGTVP	VGRVETGILK	PGDVVTFAPN	NLTT-----	--EVKSIEM	HHEALAEATP	GDNVGFNVKN	VSV-KDIRRG	FV-----	-----	-----	[312]
090_EF1a P. tricophorum FJ807243	-KIGGIGTVP	VGRVETGILK	PGDVVTFAPN	NLTT-----	--EVKSIEM	HHEALAEATP	GDNVGFNVKN	VSV-KDIRRG	FVASNSK---	N-DP-AKET-	--ADFTAQV	[329]
091_EF1a_Colacium_elongatum	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--GDFTAQV	[331]
092_EF1a_Colacium_mucronatum	-KIGGIGTVP	VGRLETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLLEAP	GDNVGFNVKN	VSL-KDIRRG	YVCSNAK---	N-DP-AKEA	--GDFTAQV	[313]
093_EF1a_Colacium_sp.	-KIGGIGTVP	VGRVETGILK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--GDFTAQV	[331]
094_EF1a_Colacium_vesiculosum	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--GDFTAQV	[317]
095_EF1a_Cryptoglena_pigra	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEAQEAVP	GDNVGFNVKN	VAV-KDIRRG	YVCSNSK---	N-DP-AKEA	--GDFTAQV	[331]
096_EF1a_Cryptoglena_sklajai	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAVP	GDNVGFNVKN	VAV-KDIRRG	FVCSNSK---	N-DP-AKEA	--ADFTAQV	[331]
097_EF1a_Discoplistis_adunca	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--ADFTAQV	[331]
098_EF1a_Euglena_clara	-KIGGIGTVP	VGRVETGILK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLLEAP	GDNVGFNVKN	VSV-KDIRRG	FVASNSK---	N-DP-AKEA	--ADFTAQV	[291]
099_EF1a_Euglena_deses	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLLEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--GDFTAQV	[331]
100_EF1a_Euglena_exilis	-KIGGIGTVP	VGRVETGILK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGFNVKN	VSV-KDIRRG	FVCSNAK---	N-DP-AKEA	--SDFTAQV	[291]
101_EF1a_Euglena_geniculata	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--ADFTAQV	[331]
102_EF1a_Euglena_gracilis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--ADFTAQV	[331]
103_EF1a E. gracilis X16890	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--ADFTAQV	[332]
104_EF1a E. gracilis FJ807238	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--ADFTAQV	[332]
105_EF1a_Euglena_granulata	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--ADFTAQV	[331]
106_EF1a_Euglena_hiemalis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--ADFTAQV	[331]
107_EF1a_Euglena_longa	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--ADFTAQV	[170]
108_EF1a_Euglena_stellata	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLLEAP	GDNVGFNVKN	VSV-KDIRRG	FVCSNAK---	N-DP-AKEA	--GDFTAQV	[291]
109_EF1a_Euglena_viridis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	FVCSNAK---	N-DP-AKEA	--SEFTAQV	[317]
110_EF1a_Eutreptia_pertyi	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--ADFTAQV	[331]
111_EF1a_Eutreptiella_pomquetensis	-KIGGIGTVP	VGRVETGV?K	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--ADFTAQV	[201]
112_EF1a_Lepocinlis_acus_var._major	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VAV-KDIRRG	YVCSNSK---	N-DP-AKEA	--ADFTAQV	[331]
113_EF1a_Lepocinlis_ovum	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLPEAVP	GDNVGFNVK?	VSV-KDIRRG	YVCSNSK---	N-DP-AKEA	--ADFTAQV	[291]
114_EF1a_Lepocinlis_spyrogyroides_clone_38	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	FVCSNAK---	N-DP-AKEA	--ADFTAQV	[102]
115_EF1a L. spyrogyroides clones_1_&_4	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VAV-KDIRRG	YVASNSK---	N-DP-AKEA	--ADFTAQV	[331]
116_EF1a_Lepocinlis_steinii	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	FVCSNAK---	N-DP-AKEA	--ADFTAQV	[291]
117_EF1a_Lepocinlis_tripteris	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--GDFTAQV	[331]
118_EF1a_Monomorphina_aenigmatica	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--ADFTAQV	[331]
119_EF1a_Monomorphina_pyrum	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VAV-KDIRRG	YVASNSK---	N-DP-AKEA	--ADFTAQV	[331]
120_EF1a_Monomorphina_sp.	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VAV-KDIRRG	YVASNAK---	N-DP-AKEA	--ADFTAQV	[331]
121_EF1a_Phacus_orbicularis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VSV-KDIRRG	FVASNAK---	N-DP-AKEA	--SDFTAQV	[317]
122_EF1a_Phacus_similis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALPEAP	GDNVGFNVKN	VSV-KDIRRG	FVCSNAK---	N-DP-AKEA	--SDFTAQV	[331]
123_EF1a_Strombomonas_acuminata	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALPEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--GDFTAQV	[332]
124_EF1a_Strombomonas_borsteniensis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSAEM	HHESLPEAVP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--ADFTAQV	[291]
125_EF1a_Strombomonas_napiformis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--GDFTAQV	[331]
126_EF1a_Strombomonas_ovalis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLPEAVP	GDNVGFNVKN	VSV-KDIRRG	YVCSNAK---	N-DP-AKEA	--GDFTAQV	[331]
127_EF1a_Strombomonas_triquetra	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALAEAP	GDNVGF----	-----	-----	-----	-----	[297]
128_EF1a_Trachelomonas_conradi	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--GDFTAQV	[331]
129_EF1a_Trachelomonas_echinata	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	YVASNAK---	N-DP-AKEA	--GDFTAQV	[331]
130_EF1a_Trachelomonas_lefevrei	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VAV-KDIRRG	YVASNAK---	N-DP-AKEA	--GDFTAQV	[331]
131_EF1a_Trachelomonas_planchonica	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHEALTEAP	GDNVGFNVKN	VAV-KDIRRG	YVASNAK---	N-DP-AKEA	--ADFTAQV	[291]
132_EF1a_Trachelomonas_similis	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	NLTT-----	--EVKSVM	HHESLAEAP	GDNVGFNVKN	VSV-KDIRRG	FVCSNAK---	N-DP-AKEA	--GDFTAQV	[291]

	340	350	360	370	380	390	400	410	420	430	440	
133_EFL1a_Trachelomonas_volvocina	-KIGGIGTVP	VGRVETGVLK	PGDVVTFAPN	ALTT-----	---EVKSIVEM	HHEALSEAVP	GDNVGFNVKN	VSV-KDIRRG	FVCSNAK---	N-DP-AKEA-	---GDFTAQV	[331]
134_EFLenvironmental_sample_EU810232	-AIKGTGDVL	TGRIEQGSVK	PGDEVVFLPT	HTAANPCA--	--GKVFSIEM	HHKTLDAALP	GHNVGMNIKG	LPKGNMPRTG	DIMILKS---	D-TS-LKPC-	---KSFAQV	[362]
135_EFLCreolimax_fragrantissima	-KIKGVGDVL	TGRVEQGTVK	PGDECIFLPT	HTPSTGCS--	--GKVFTVEM	HHKQCEAAYA	GDNVGLNIKG	LNAKNMPRAG	DVMILKS---	D-TT-LGRC-	-----	[331]
136_EFLSpheeroforma_arctica	-KIKGVGDVL	TGRVEQGTVR	PGDECVFLLP	HTPSTGCT--	--GKVFTVEM	HHKQVEAAYA	GDNVGLSIKG	LNAKNMPRAG	DVMILKS---	D-NT-LGRC-	---KNFTAQV	[362]
137_EFLMonosiga_brevicollis	-KIKGVGDVL	TGRVEQGTVK	PGDEVVFLPT	HTSSTACT--	--GKVFTVEM	HHKGVVEAAMT	GDNVGLNIKG	LKNDNMPRVG	DVMILKS---	D-DS-LGRV-	---KSFTVQV	[362]
138_EFLXenopus_laevis	-KIKGVGDVL	TGRVEQGTVK	PGDEVIFLPT	HTPSTACT--	--GKVFSVEM	HHKQVESAGP	GDNVGMNVKN	LNAKNMPRTG	DVMILKS---	D-QS-LGRC-	---KNFTVQV	[362]
139_EFLSpizellomyces_punctatus	-KIKGVGDVL	AGRVEQGTVL	PNKDVFVFP	HTAANPCG--	--GKVFTVEM	HHKQVEKGLA	GDNIGMNIKG	LKNNMPRAG	DIMIYKE---	D-KS-LKPV-	---KAFTAQV	[364]
140_EFLGaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[168]
141_EFLAllomyces_macrogynus	-KIKGVGDVL	AGRVEQGIPT	TNKDVVFLPT	HTTATPCA--	--GKVFTVEM	HHKQVEKGET	GDNVGMNIKG	LKNNMPRVG	DVMIYKE---	D-ST-LKAV-	---KSFTAQV	[364]
142_EFLBlastocladiella_emersonii	-KIKGVGDVL	AGRVEQGIPT	PNKDVFVFLPT	HTTATPCA--	--GKVFTVEM	HHKQVEKGET	GDNVGMNIKG	LKNNMPRVG	DVMIYKE---	D-AT-LKAV-	---KSFTAQI	[364]
143_EFLCatenophylctis_sp.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[168]
144_EFLCoelomomyces_stegomyiae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[168]
145_EFLEntomophthora_muscae	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[167]
146_EFLBasidiobolus_ranarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[167]
147_EFLConidiobolus_coronatus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[167]
148_EFLTetraselmis_tetrathele	-KIKGVGDVL	AGRVEQGVK	PNDEVVFMPT	HTKANSCT--	--GKVFTVEM	HHKRVAAAGP	GDNVGMNIKG	LEKNNMPRSG	DVMIHKS---	D-TT-LGHV-	---EDFTAQI	[345]
149_EFLMesostigma_viride	-KIKGVGDVL	AGRVEQGIPT	PGEELV----	-----	-----	-----	-----	-----	-----	-----	-----	[293]
150_EFLOstreococcus_tauri	-KIKGVGDVL	AGRVEQGLVK	PGDEVVFLPT	HTSSNPCG--	--GKVFTVEM	HHKRVAAAP	GDNVGMNIKG	LDKLNMPRTG	DVMIYK--	D-TS-LAPC-	---KNFTAQV	[365]
151_EFLNephrolelmis_olivacea	-KIKGVGDVL	AGRVEQGTVC	PGDEVIFLPT	HTASNPCV--	--GKVFSVEM	HHKSVAEAKP	GDNVGM----	-----	-----	-----	-----	[319]
152_EFLMicromonas_pusilla	-KIKGVGDVL	AGRVEQGVVR	PGEVIFLPT	HTVSNPCV--	--GKVFTVEM	HHKRVAAEAP	GDNVGMNIKG	LKNNMPRSG	DVMIYK--	D-TT-LKAC-	---KNFNAQI	[348]
153_EFLParachlorella_kessleri	-KIKGVGDVL	AGRVEQGIPT	PGDEVIFLPT	HTAANPCI--	--GKVFTVEM	HHKRVDAKAG	GDNVGM----	-----	-----	-----	-----	[314]
154_EFLHelicosporidium_sp.	-KIKGVGDVL	AGRVEQGIPT	PNDEVVFMPT	HTAANQCT--	--GKVFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRTG	DVMIYK--	D-AT-LAPV-	---KDFTAQI	[369]
155_EFLGonioloma_pectinale	-KIKGVGDVL	AGRVEQGVVK	PGDEVVFLPT	HTTANPCS--	--GKVFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRTG	DVMILKS---	D-ST-LKLV-	---QSFTAQI	[363]
156_EFLChlamydomonas_reinhardtii	-KIKGVGDVL	AGRVEQGVVK	PGDEVIFLPT	HTTANPCS--	--GKVFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRTG	DVMILKS---	D-QT-LKIV-	---KDFTAQI	[363]
157_EFLScenedesmus_obliquus	-KIKGVGDVL	AGRVEQGVVK	PGDEVIFLPT	HTTANPCV--	--GKVFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRTG	DVMILKS---	D-QT-LKPC-	---KDFTAQI	[363]
158_EFL"Oryza_sativa"	-KIKGVGDVL	AGRVEQGVVK	PGDEVIFLPT	HTAANPCV--	--GKVFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRTG	DVMILKS---	D-QT-LKPC-	---QNFTAQI	[363]
159_EFLChlorococcum_sp.	-KIKGVGDVL	AGRVEQGIPT	PGEELVFMPT	HTAANKCE--	--GKVFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRAG	DVMIHKT---	D-GT-LKHV-	---ANFTAQI	[346]
160_EFLOchlochaete_hystrix	-KIKGVGDVL	AGRVEQGVVK	PGDEVVFMPT	HTPQNPCT--	--GKVFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRTG	DVMILKS---	D-TT-LKQV-	---KDFTAQI	[349]
161_EFLUlva_intestinalis	-KIKGVGDVL	AGRVEQGIPT	PNDEVVFMPT	HTAANKCE--	--GKVFTVEM	HHKRVDSAGP	GDNVGMNIKG	LKNNMPRSG	DVMIMKS---	D-PS-LKHV-	---EDFSAQI	[346]
162_EFLUrospora_sp.	-KIKGVGDVL	AGRVEQGVVK	PNDEVVFMPT	HTAANKCE--	--GKVFSVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRAG	DVMIMKS---	D-AS-LKHV-	---EDFSAQI	[297]
163_EFLGymnochlora_stellata_ppc	-RIKGVGDVL	AGRVEQGVVK	NEADVIFLPT	HVPNHPCN--	--GKVFSIEM	HHKRVDAKAG	GDNVGMNIKG	FPKDNPPRSG	DVMILSG---	D-KS-LQST-	---GEFTAQV	[375]
164_EFLLotharella_amoebiformis_ppc	-RIKGVGDVL	AGRVEQGVVK	NEADVIFLPT	HVPNHPCN--	--GKVFSIEM	HHKRVDAKAG	GDNVGMNIKG	FPKDNPPRSG	DVMILSG---	D-KS-LQST-	---GEFTAQV	[290]
165_EFLChlorarachnion_reptans_ppc	-RIKGVGDVL	AGRVEQGVVK	NEADVIFLPT	HTPNNPCS--	--GKVFSIEM	HHKRVDAKAG	GDNVGMNIKG	FPKDNPPRSG	DVMILSG---	D-DS-LHAT-	---GEFTAQV	[299]
166_EFLChondrus_crispus	-KIKGVGDVL	TGRVEQGTVC	PGDEVIFLPT	HTTSTPCS--	--GKVFSVEM	HHKSVPAAEP	GDNVGLNVKG	VEKLNMPRVG	DVMILTK---	D-DS-LKRC-	---ATFTAQV	[362]
167_EFLGracilaria_changii	-KIKGVGDVL	TGRVEQGVVK	PGDEVIFLPT	HTSSTPCS--	--GKVFSVEM	HHKSVPAAEP	GDNVGLNVKG	VEKLNMPRVG	DVMILTK---	D-DS-LKRC-	---ATFTAQV	[362]
168_EFLCryptomonas_ovata	-KIKGVGDVL	AGRVEQGLVN	PGDEVVFMPT	HASGTPCA--	--GKIFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRSG	DVMYVYK--	D-GT-LKGT-	---KSFTAQI	[348]
169_EFLGuillardia_theta	-KIKGVGDVL	AGRVEQGLVN	PGDEVVFMPT	HTPGTPCE--	--GKIFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRSG	DVMYVYK--	D-GT-LKGT-	---KSFTAQI	[363]
170_EFLRhodomonas_salina	-KIKGVGDVL	AGRVEQGLVN	PGDEVVFMPT	HTTGTPCA--	--GKIFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRSG	DVMYVYK--	D-GT-LKGT-	---KSFTAQI	[362]
171_EFLPavlova_lutheri	-KIKGVGDVL	AGRVEQGIPT	PGDEVIFLPT	HTTVNPCV--	--GKVFTVEM	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRSG	DVMYVYK--	D-GT-LKAA-	---ENFNAQI	[362]
172_EFLChrysochromulina_sp.	-KIKGVGDVL	AGRVEQGLVK	PNDEVVFMPT	HTAANACG--	--GKVFTVEM	HHKRVDAKAG	GDNVGLNIKG	LKNNMPRVG	DVMYVYK--	D-SS-LDRC-	---GQFTAQI	[302]
173_EFLIsochrysis_galbana	-KIKGVGDVL	AGRVEQGLVK	PNDEVVFMPT	HTAANACA--	--GKVFTVEM	HHKRVDAKAG	GDNVGLNIKG	LKNNMPRVG	DVMYVYK--	D-KS-LDQV-	---GQFDAQV	[363]
174_EFLEmiliana_huxleyi	-KIKGVGDVL	AGRVEQGVIT	PNDEVVFMPT	HTAANACA--	--GKVFTVEM	HHKRVDAKAG	GDNVGLNIKG	LKNNMPRVG	DVMYVYK--	D-KT-LTQC-	---GEFDAQI	[365]
175_EFLAchnanthes_kuwaitensis	-NIKGVGQII	CGTVEQGTIR	PGDVIGMVPS	NLKG-----	--KKIFSIEQ	HHKTLDSAGP	GNSVGLSIKG	IAKDEKVQPG	DIYIYK---	E-GD-LKPI-	---KSFTAMV	[332]
176_EFLAsterionella_glacialis	-NIKGVGQII	CGTVEQGTIR	PGDVIGMVPS	NLKG-----	--KKIFSIEQ	HHKTLDSAGP	GNSVGLSIKG	IAKDEKVQPG	DIYIYK---	E-GD-LKPI-	---KSFTAMV	[332]
177_EFLDetonula_confervacea	-NIKGVGQII	CGTVEQGTIR	PGDVIGMVPS	NLKG-----	--KKIFSIEQ	HHKTLDSAGP	GNSVGLSIKG	IAKDEKVQPG	DIYIYK---	E-GD-LKST-	---KSFTAMV	[332]
178_EFLThalassiosira_pseudonana	-NIKGVGQII	CGTVEQGTIR	PGDVIGMVPS	NLKG-----	--KKIFSIEQ	HHKTLDSAGP	GNSVGLSIKG	IAKDEKVQPG	DIYIYK---	E-GD-LKPI-	---KSFTAMV	[353]
179_EFLPerkinsus_marinus	-KIKGVGDVL	TGRVEQGIPT	PNDEVVFMPT	HTPSTACT--	--GKVFTVEM	HHKRVDAKAG	GDNVGLNVKG	LKNNMPRSG	DVMYVYK--	D-TT-LKHV-	---QSFTAQV	[368]
180_EFLOxyrrhis_marina	-KIKGVGDVL	AGRVEQGIPT	PGDEVVFLPT	HTPSNACS--	--GKVFTVEM	HHKRVDAKAG	GDNVGLNIKG	LKNNMPRAG	DVMYVYK--	D-TT-LSAA-	---GSFTAQV	[366]
181_EFLKarlodinium_micrum	-KIKGVGDVL	AGRVEQGIPT	PGDEVVFLPT	HTTSPNCT--	--GKVFTVEM	HHKRVDAKAG	GDNVGLNIKG	LKNNMPRSG	DVMYVYK--	D-TT-LGQT-	---KEFDAQI	[364]
182_EFLAmphidinium_carterae	-KIKGVGDVL	AGRVEQGIPT	PGDEVVFLPT	HTASNPCT--	--GKVFTVEM	HHKRVDAKAG	GDNVGLNIKG	LKNNMPRSG	DVMYVYK--	D-ST-LGQT-	---KAFDAQI	[364]
183_EFLHeterocapsa_triquetra	-KIKGVGDVL	AGRVEQGIPT	PGDEVVFLPT	HTASNPCT--	--GKVFTVEM	HHKRVDAKAG	GDNVGLNIKG	LKNNMPRSG	DVMYVYK--	D-TT-LGQT-	---KEFDAQI	[364]
184_EFLPlanoglabratella_operculis	-KIKGVGDVL	TGRVEQGTVC	PGDVVFRFPT	GGC-----	--GKAFSIEQ	HHKRVDAKAG	GDNVGMNVKG	LKNNMPRSG	DVMYVYK--	D-PPPKQA-	---SAFTALV	[309]
185_EFLReticulomyxa_filosa	-NIKGVGDVL	TGRVEQGTVC	PGDVVFRFPT	NAT-----	--GKAFSIEQ	HHKRVDAKAG	GDNVGMNVKG	LKNNMPRSG	DVMYVYK--	D-PPPKQA-	---SAFTALV	[333]
186_EFLLotharella_amoebiformis	-KMK-AGTIV	TGRIEQGVLE	KEVKTKTGVS	GNPFRFVPSG	LEAKVFSIEA	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRSG	DVMYVYK--	D-PPPKQA-	---SAFTALV	[389]
187_EFLGymnochlora_stellata	-KMS-AGTIV	TGRIEQGVLE	KEVKTKTGTT	GTPVKFVPSG	LEAKVFSIEA	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRSG	DVMYVYK--	D-PPPKQA-	---SAFTALV	[389]
188_EFLChlorarachnion_reptans	-KMS-AGTIV	TGRIEQGVLE	KEVKTKTGTT	GTPVKFVPSG	LEAKVFSIEA	HHKRVDAKAG	GDNVGMNIKG	LKNNMPRSG	DVMYVYK--	D-PPPKQA-	---SAFTALV	[389]
189_EFLDiplonema_ambulator	-KIKGVGDVL	TGRVEQGVVK	PGDEVVFLPT	HTAANACT--	--GKVFTCEM	HHKRVDAKAG	GDNVGLNVKG	LKNNMPRSG	DVMYVYK--	D-GS-LKEA-	---GTFNAQV	[230]
190_EFLDiplonema_papillatum	-KIKGVGDVL	TGRVEQGVVK	PGDEVVFLPT	HTAANACT--	--GKVFTCEM	HHKRVDAKAG	GDNVGLNVKG	LKNNMPRSG	DVMYVYK--	D-TS-LKEA-	---GTFNAQV	[354]
191_EFLRhynchopus_euleeides	-KIKGVGDVL	TGRVEQGVVK	PGDEVVFLPT	HTAANACT--	--GKVFTCEM	HHKRVDAKAG	GDNVGLNVKG	LKNNMPRSG	DVMYVYK--	D-AT-LKEA-	---GTFNAQV	[235]
192_EFLPetalomonas_cantuscygni	-KIKGVGDVL	TGRVEQGIPT	PGDEVVFLPT	HTASNPCT--	--GKVFSIEM	HHKRVDAKAG	GDNVGLNVKG	LKNNMPRSG	DVMYVYK--	D-PS-LKSA-	---GQFT---	[308]
193_EFLTrypanoplasma_borrel	-SIKGVGDVL	TMRIEQGVLE	KGDEVIFLPT	HTSSTPCQ--	--GKVFSIEM	HHKRVDAKAG	GDNVGLNVKG	LKNNMPRSG	DVMYVYK--	D-PS-LKSA-	---GQFT---	[339]
194_EFLNeobodo_salien	TSIKGIECVV	TSRLEQGTIV	KGDEVVFLPT	HTASNPCT--	--GKVFSIEM	HHKRVDAKAG	GDNVGMNVKG	LKNNMPRSG	DVMYVYK--	D-TD-FKAC-	---ESFVQA	[353]

	450	460	470	480	490	500	510	520	530	540	550	
001_aEF1a_Desulfurococcus_mobilis	MVMV-HPTAI	AV-----	--GYTPVIH	VHTASVACRI	TEII--AKI	DPR--TGKE	I---EKNPH	FLKQGDIAI-	VKFKPIKPLV	VEKYSDFQGL	GRFAMRDMGK	[419]
002_aEF1a_Sulfolobus_solfataricus	IVVW-HPTAL	AN-----	---GYTPVIH	VHTASVACRV	SELV---SKL	DPR--TGQE	A---EKNPH	FLKQGDVAI-	VKFKPIKPLC	VEKYNEFPPL	GRFAMRDMGK	[417]
003_aEF1a_Thermoplasma_acidophilum	IVLN-HPSVI	AP-----	---GYKPVFH	VHTAQVACRI	DEIV---KTL	NPK---DGTG	L----KEKPD	FIKNGDVAI-	VKVIPDKPLV	IEKVSEIPQL	GRFAVRDMGQ	[414]
004_aEF1a_Halobacterium_sp.	VVMQ-HPSVI	TA-----	---GYTPVFH	AHTAQVACTI	ESID---KKM	DPA---SGET	Q----EENPD	FIQSGDAAV-	VTVRPQKPLS	LEPSSSEIPEL	GSFAVRDMGQ	[408]
005_aEF1a_Methanocaldococcus_jannaschii	VVLQ-HPTAI	TV-----	---GYTPVVFH	AHTAQVACTF	TELL---KKL	DPR--TGQV	I---EENPD	FLKTGDAAI-	VKIKPTKPMV	IENVREIPQL	GRFAIRDMMG	[413]
006_EF1a_Acanthamoeba_culbertsoni	IILN-HPGQI	HA-----	---GYAPVLD	CHTAHIACKF	KELL---EKV	DRR--SGKK	L----EDSPK	AVKSGDAAM-	VLLVPSKPLC	VETFTYEYPL	GRFAVRDMRQ	[408]
007_EF1a_Neoparamoeba_branchipila	IVLN-HPGEI	RS-----	---GYTPVLD	CHTAHIACKF	AEIT---EKI	DRR--SGKT	V----EKEPK	AVKSGDAAI-	VRLTPSKPMC	VETFKDYAPL	GRFAVRD---	[400]
008_EF1a_Planoprotostelium_aurantium	IILN-HPGEI	HN-----	---GYAPVLD	CHTAHIACKF	AEIT---SKI	DRR--SGKE	L----EALPK	SVKSGDAAM-	VKLVPQKPMC	VETFTYEYPL	GRFAVRDMR-	[401]
009_EF1a_Physarum_polycephalum	IILN-HPGTI	HA-----	---GYAPVLD	CHTAHIACKF	TEIL---SKV	DRR--TGQA	V----EEVPK	NIKNGDAAI-	VKLTPSKPMC	VESFTDFPPL	GRFAVRDMR-	[401]
010_EF1a_Amoebidium_parasiticum												[289]
011_EF1a_Capsaspora_owczaraki	IVLN-HPGQI	SN-----	---GYAPVLD	CHTAHIACKF	QDIK---EKC	DRR--SGKK	L----EDAPK	FVKSGDAAI-	VELLPTKPMC	VEAFSDYPPPL	GRFAVRDMRQ	[429]
012_EF1a_Monosiga_ovata	IILN-HPGQI	SN-----	---GYAPVLD	CHTAHIACKF	NEIK---EKI	DRR--SGKK	I----EDFPK	AVKSGDAAI-	VEMIPKKPMC	VEPFSEYPPPL	GRFAVRDMRQ	[422]
013_EF1a_Salpingoeca_amphoridium	IIMN-HPGQI	SN-----	---GYAPVLD	CHTAHIACKF	HEIT---AKL	DRR--TGKA	T----EENPK	FVKNGDAAL-	VDMPQKPMC	VETFAEYPPPL	GRFAVRDMRQ	[423]
014_EF1a_Ceanorhabditis_elegans	IIMN-HPGQI	SN-----	---GYTPVLD	CHTAHIACKF	NELK---EKV	DRR--TGKK	V----EDFPK	FLKSGDAGI-	VELIPTKPLC	VESFTDYAPL	GRFAVRDMRQ	[431]
015_EF1a_Drosophila_melanogaster	IVLN-HPGQI	AN-----	---GYTPVLD	CHTAHIACKF	AEIK---EKV	DRR--SGKT	T----EENPK	FIKSGDAAI-	VNLVPSKPLC	VEAFQEFPPPL	GRFAVRDMRQ	[431]
016_EF1a_Ornithorhynchus_anatinus	IILN-HPGQI	SA-----	---GYSPVID	CHTAHIACKF	AELK---EKI	DRR--SGKK	L----EDNPK	SLKSGDAAI-	VEMIPGKPMC	VESFSQYPPPL	GRFAVRDMRQ	[431]
017_EF1a_Xenopus_laevis	IILN-HPGQI	SA-----	---GYAPVLD	CHTAHIACKF	AELK---QKI	DRR--SGKK	L----EDDPK	FLKSGDAAI-	VEMIPGKPMC	VETFSQYPPPL	GRFAVRDMRQ	[431]
018_EF1a_Saccharomyces_cerevisiae	IVLN-HPGQI	SA-----	---GYSPVLD	CHTAHIACRF	DELL---EKN	DRR--SGKK	L----EDHPK	FLKSGDAAL-	VKFVPSKPMC	VEAFSEYPPPL	GRFAVRDMRQ	[429]
019_EF1a_Neuropora_crassa	IVLN-HPGQV	GA-----	---GYAPVLD	CHTAHIACKF	AELL---EKI	DRR--TGKA	V----EASPK	FIKSGDAAI-	VKMIPSKPMC	VEAFTDYPPPL	GRFAVRDMRQ	[430]
020_EF1a_Cryptococcus_neoformans	IVLN-HPGQI	GA-----	---GYTPVLD	CHTAHIACKF	AELI---EKI	DRR--TGKV	M----EAAPK	FVKSGDAAI-	VKLVAQKPLC	VETYADYPPPL	GRFAVRDMRQ	[429]
021_EF1a_Schizophyllum_comune	IVLN-HPGQI	GA-----	---GYAPVLD	CHTAHIACKF	AELL---EKI	DRR--TGKS	L----EASPK	FVKSGDACT-	VKLVPKPMC	VESFYEYPPPL	GRFAVRDMRQ	[429]
022_EF1a_Monoblepharis_insignis	IVMN-HPGQI	SN-----	---GYSPVLD	CHTAHIACRF	NNIL---QKI	DRR--SGKV	L----EENPK	FIKSGDAAM-	VEMIPTKPMC	VESFNEYPPPL	GRFAVRDMR-	[390]
023_EF1a_Rhizophlyctis_harderi	IVLN-HPGQI	SA-----	---GYSPVLD	CHTAHIACKF	AELL---EKL	DRR--TSKK	L----EDNPK	FIKSGDAAI-	VKMVPSKPMC	VESFTEYPPPL	GRFAVRDMR-	[390]
024_EF1a_Harpochytrium_sp.	IILN-HPGQI	GA-----	---GYAPVLD	CHTAHIACKF	AELV---EKI	DRR--SGKK	M----EDAPK	FVKSGDACT-	AKMVPSKPMC	GESF-----		[359]
025_EF1a_Absidia_glauca	IVLN-HPGQI	GA-----	---GYAPVLD	CHTAHIACKF	AELL---EKI	DRR--SGKK	L----EDAPK	FVKSGDAAI-	VKMIPSKPMC	VEAFTDYPPPL	GRFAVRDMRQ	[429]
026_EF1a_Smittium_simulii	IVLN-HPGQI	SN-----	---GYAPVLD	CHTAHISCKF	TEII---EKM	DRR--SGKT	L----EAAPK	FIKSGDAAI-	VVLTPSKPMC	VESFQYPPPL	GRFAVRDMRQ	[420]
027_EF1a_Chara_australis	IIMN-HPGQI	GN-----	---GYAPVLD	CHTSHIAVKF	AELV---VKI	DRR--TGKE	I----EKEPK	FLKNGDAGF-	IKMIPTKPMC	VETFAEYPPPL	GRFAVRDMRQ	[402]
028_EF1a_Physcomitrella_patens	IIMN-HPGQI	GN-----	---GYAPVLD	CHTSHIAVKF	AELV---TKV	DRR--SGKE	L----EKEPK	FLKNGDAGF-	VKMIPTKPMI	VETFAEYPPPL	GRFAVRDMRQ	[419]
029_EF1a_Acetabularia_acetabulum	IIMN-HPGQI	GN-----	---GYQPVLD	CHTAHIAVKF	KDII---KKV	DRR--SGKV	V----EESPK	FIKNGDAAF-	VIMEPSKPMV	VESFTEYPPPL	GRFAVRDMRQ	[402]
030_EF1a_Chaetomorpha_coliformis	IIMN-HPGQI	SN-----	---GYTPVLD	CHTSHIAVKF	KEIQ---QKV	DRR--TGKV	T----EENPK	FIKNGDAGF-	VLMPEKPMC	VEPFTYEYAPL	GRFAVRDMRQ	[352]
031_EF1a_Arabidopsis_thaliana	IIMN-HPGQI	GN-----	---GYAPVLD	CHTSHIAVKF	SEIL---TKI	DRR--SGKE	I----EKEPK	FLKNGDAGM-	VKMIPTKPMV	VETFSYPPPL	GRFAVRDMRQ	[419]
032_EF1a_Oryza_sativa	IIMN-HPGQI	GN-----	---GYAPVLD	CHTSHIAVKF	AELV---TKI	DRR--SGKE	L----EKEPK	FLKNGDAGM-	VKMIPTKPMV	VETFSYPPPL	GRFAVRDMRQ	[419]
033_EF1a_Cyanidioschyzon_merolae	IILN-HPGEI	RA-----	---GYAPVLD	CHTAHIACRF	AELL---EKI	DRR--TGKK	I----EENPE	KVKAGDACH-	VRMIPSKPMC	VETFSYPPPL	GRFAVRDMRQ	[423]
034_EF1a_Porphyrha_jezoensis	IILN-HPGEI	HA-----	---GYAPVLD	CHTAHIACKF	SELI---LKM	DRR--SGKK	L----EDTPK	MIKSGDAAI-	VKMVASKPMC	VEAFTQYPPPL	GRFAVRDMRQ	[423]
035_EF1a_Plasmodium_falciparum	IILN-HPGEI	KN-----	---GYTPVLD	CHTSHISCKF	LNDI---SKI	DKR--SGKV	V----EENPK	AIKSGDSAL-	VSLEPKPMV	VETFTYEYPL	GRFAIRDMMQ	[417]
036_EF1a_Cryptosporidium_parvum	IVLN-HPGEI	KN-----	---GYSPVLD	CHTAHISCKF	QTIH---AKM	DKR--SGKV	L----EENPK	LIKSGDAAI-	VVMQPLKPLC	VEAFTDYPPPL	GRFAVRDMKQ	[415]
037_EF1a_Tetrahymena_pyriformis	IIMN-HPGQI	QA-----	---GYTPVLD	CHTAHIACKF	ETIH---DKI	DRR--TGKS	Q----EENPK	FIKNGDAAI-	VTLIPTKALC	VEVFQYPPPL	GRYAVRDMKQ	[420]
038_EF1a_Stentor_coeruleus	IVLN-HPGQI	SA-----	---GYTPVLD	CHTAHIACRF	NIHQ---KKI	DRR--TVAS	L----EDELK	FVKSGDAAFD	VLGSPSKPMC	VETFTYEYPL	GRFAVRDMRQ	[405]
039_EF1a_Phytophthora_parasitica	IVLN-HPGQI	GN-----	---GYSPVLD	CHTAHVACQI	KEIT---EKM	DRR--SGKV	L----ETAPK	FVKSGDACH-	VILEPSKPMI	VESFQYPPPL	GRFAVRDMRQ	[419]
040_EF1a_Mallomonas_rasilis	IIMS-HPGQI	QN-----	---GYTPVLD	CHTSHIATKF	KNID---EKM	DRR--TGKS	M----EDNPK	FVKSGDACH-	VTMEPTKPMV	VEAFSDYPPPL	GRFAVRDMRQ	[406]
041_EF1a_Phaeodactylum_tricornutum	IVMN-HPGQI	SN-----	---GYSPVLD	CHTAHVACQI	ALIK---EKM	DRR--SGKV	L----EQNPK	FVKTGDACT-	VDLEPTKPLC	VESFTDFPPL	GRFAVRDMRQ	[419]
042_EF1a_Thalassiosira_pseudonana	IIMN-HPGKI	FV-----	---GYTPVLD	CHTAHVACQI	TELK---QRM	DRT--NGQV	L----EDNPK	FVKAGDMCL-	ATLVPTKPLC	VETFTFEAPL	GRFAVRDMKR	[419]
043_EF1a_Malawimonas_jakobiformis	IILN-HPGQI	GA-----	---GYTPVLD	CHTAHIACKF	NEIV---AKV	DRR--TGKA	S----EEEPK	FIKSGDAAI-	VKLVPKPMC	VESFTEYPPPL	GRFAVRDMR-	[395]
044_EF1a_Malawimonas_californiana	IVLN-HPGQI	GA-----	---GYTPVLD	CHTAHIACKF	AELL---QKV	DRR--TGKV	L----EEEPK	FVKSGDAAM-	VKLIPSKPMC	VETFSYPSPL	GRFAVRDMRQ	[419]
045_EF1a_Trimastix_pyriformis	IIMN-HPGEI	HS-----	---NYTPVLD	CHTSHIACKF	EELL---TKI	DRR--TGRA	T----EEFPK	MVKTGDSAL-	VQIVPTKPMV	VEPFTYEYPL	GRFAVRDMR-	[404]
046_EF1a_Dinenympha_exilis	IIVMA-HPGQI	QN-----	---GYTPVLD	CHTSHIACKF	KEIQ---SKI	DRR--TNKV	Q----EENPK	FIKTGDSAI-	VEMQPSKPMV	VETFTYEYPL	GRFAVR---	[395]
047_EF1a_Monocercomonoides_sp.	IIVMS-HPGQI	SN-----	---GYTPVLD	CHTAHIACKF	KEIT---AKI	DRR--TNKV	Q----EENPK	FIKTGDSAL-	VELVPSKPMV	VEAFTYEYPL	GRFAVRDMRA	[419]
048_EF1a_Streblospastrix_strix	IIVMQ-HPGQI	QN-----	---GYTPVLD	CHTSHIACKF	KEIT---SKI	DRR--TNKE	Q----EANPK	SIKTGDSAI-	VTMVPSKPMV	VEKFVEYPPPL	GRFAVR---	[398]
049_EF1a_Histomonas_meleagridis	IISN-HPGKI	HA-----	---GYQPVFD	CHTAHIACKF	DKLV---QRI	DRR--HGKK	V----TENPE	WIKQDDAAV-	VVVVPSKPLV	VETFKYEAAL	GRFAVRDMKQ	[419]
050_EF1a_Trichomonas_vaginalis	IISN-HPGKI	HA-----	---GYQPVFD	CHTAHIACKF	DKLI---QRI	DRR--HGKK	A----TENPE	YIQKDDAAI-	VEVVPSKPLV	VESFQYPPAL	GRFAIRDMMQ	[419]
051_EF1a_Hexamita_inflata	IIMN-HPKKI	NP-----	---GYTPVLD	CHTSHLACKF	DKFL---AKL	NSR--TFKV	E----IENPT	EAVRGECVL-	MQIVPTKPLC	VESFQYPPAL	GRFAVRDMKR	[403]
052_EF1a_Spiroplasma_vortens	IIMN-HPKKI	QP-----	---GYTPVLD	CHTSHLACKF	TKFL---QKL	NSR--TLKP	E----QENPT	EARGECII-	AKITPTKEFS	AESYELYPAL	GRFAVR---	[400]
053_EF1a_Andalucia_incarcerata	IILN-HPGQI	GA-----	---GYCPVLD	CHTAHIACKF	AELT---KKI	DRR--SGKE	L----EANPK	FVKSGDAAI-	VKLVPQKPMC	VEAFTYEYPL	GRFAVRDMR-	[401]
054_EF1a_Seculamonas_ecuadoriensis	IVLN-HPGQI	GQ-----	---GYSPVLD	CHTSHIACRF	AELV---QKI	DRR--TGKV	M----EENPK	AIKSGEAAI-	VKLVPKPMC	VETFAEYPPPL	GRFAVRDMRQ	[419]
055_EF1a_Reclinomonas_americana	IILN-HPGQI	AA-----	---GYTPVLD	CHTAHIACRF	NELL---QKI	DRR--TGKE	L----EATPK	FVKSGDAVI-	AQLVPSKPLC	VEKFSEYAPL	GRFAVRDMRQ	[419]
056_EF1a_Histiona_aroidea	IILN-HPGQI	GA-----	---GYTPVLD	CHTAHIACRF	AELL---QKI	DRR--TRKE	L----EASPK	FVKSGDAVI-	AKMEPTKPMC	VEKFSEYAPL	GRFAVRDMRQ	[419]
057_EF1a_Jakoba_libera	IILN-HPGQI	NA-----	---GYAPVLD	CHTAHIACRF	AELI---TKI	DRR--TGKE	L----EASPK	FVKSGDVCI-	AQLVPQKPMC	VEKFSADVCP	GPFPVPPMQR	[419]
058_EF1a_Stachyamoeba_lipophora	IIMN-HPGEI	RN-----	---GYSPVID	CHTAHIACKF	EAIT---EKL	DKR--TGKA	V----EENPA	FVKKGDAAT-	VRLVPQKPLC	VEDFKTYPPPL	GRFAVRDMRQ	[419]
059_EF1a_Naegleria_gruberi	IIMN-HPGEI	RN-----	---GYTPVLD	CHTSHIACKF	ETII---EKI	DRR--TGKT	Q----EANPE	KIKKGDSAI-	VKIVPTKPMC	VEGFVEYPPPL	GRFAVRDMK-	[402]
060_EF1a_Naegleria_andersoni	IIMN-HPGEI	RA-----	---GYTPVLD	CHTSHIACKF	DKLI---DKI	DRR--TGKS	Q----EANPE	KIKKGDSAI-	VQIVPTKPMC	VEGFVEYPPPL	GRFAVR---	[405]
061_EF1a_Acrasis_rosea	IILN-HPGEI	QN-----	---GYAPVLD	CHTAHIACKF	QEIL---TKI	DKR--TGQE	M----EKLPK	SLKSGQAGI-	VNLVPQKPMC	VEEYAYPPPL	GRFAVRDMR-	[401]
062_EF1a_Sawyeria_marylandensis	IILN-HPGKI	CK-----	---GYQPVLD	CHTAHIACKF	EELI---ALV	DKR--NGQE	I----EKSPA	SLKKGDAI-	VKMVPSKSL	IESYAEYPPPL	GRFAVRDMKQ	[419]
063_EF1a_Bodo_saltans	IILN-HPGQV	GN-----	---GYSPVLD	CHTSHIACKF	AMIE---SKI	DRR--SGKE	V----EKEPK	AIKSGDAAI-				[369]
064_EF1a_Rhynchobodo_sp.	IVLN-HPGQI	GN-----	---GYAPVLD	CHTSHIACKF	QEIE---SKI	DRR--SGKE	L----EKEPK	AIKTGDAAI-				[367]
065_EF1a_Rhynchomonas_nasuta	IVLN-HPGQI	SN-----	---GYAPVLD	CHTSHIACKF	AETE---SKI	DRR--SGKV	I----EEEPK	FIKMGDAAM-	VRMIPTKPLC	VESFTD----		[404]
066_EF1a_Dimastigella_trypaniformis	IIVMN-HPGQI	SN-----	---GYAPVLD	CHTAHIACKF	TELE---AKI	DRR--TGKE	S----EAAPK	SIKNGDAAI-	I-----			[375]

	450	460	470	480	490	500	510	520	530	540	550	
067_EF1a_Perkinsiella_amoebae												[305]
068_EF1a_Leishmania_major	IVLN-HPGQI	SN-----	--GYAPVLD	CHTSHIACRF	AEIE---SKI	DRR---SGKE	L---EKNPK	AIKSGDAAI-	VKMVPQKPMC	VEVFNDYAPL	GRFAVRDMRQ	[419]
069_EF1a_Leishmania_infantum	IVLN-HPGQI	SN-----	--GYAPVLD	CHTSHIACRF	AEIE---SKI	DRR---SGKE	L---EKNPK	AIKSGDAAI-	VKMVPQKPMC	VEVFNDYAPL	GRFAVRDMRQ	[419]
070_EF1a_Leishmania_braziliensis	IVLN-HPGQI	SN-----	--GYAPVLD	CHTSHIACRF	ADIE---SKI	DRR---SGKE	L---EKNPK	AIKSGDAAI-	VKMVPQKPMC	VEVFNDYAPL	GRFAVRDMRQ	[419]
071_EF1a_Leishmania_tarentolae	IVLN-HPGQI	SN-----	-G-----									[344]
072_EF1a "L. tarentolae" FJ807257												[305]
073_EF1a_Leptomonas_costaricensis												[198]
074_EF1a_Leptomonas_bifurcata	IVLN-HPGQI	SN-----	--GYAPVLD	CHTAHIA---								[277]
075_EF1a_Leptomonas_podlipaevi	IVLN-HPGQI	SN-----	--GYAPVLD	CHTAH---								[271]
076_EF1a_Herpetomonas_muscarum	IILN-HPGQI	SN-----	--GYAPVLD	CHTSHIACRF	ADIE---SKI	DRR---SGKV	L---EEHPK	FKVSGDAAM-	VKMVPQKPMC	VEVFNDYAPL	G-----	[381]
077_EF1a_Herpetomonas_pestoi												[257]
078_EF1a_Blastocrithidia_culicis												[255]
079_EF1a_Trypanosoma_brucei	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	AEIE---SKI	DRR---SGKE	L---EKAPK	SIKSGDAAI-	VRMVPQKPMC	VEVFNDYAPL	GRFAVRDMRQ	[419]
080_EF1a_T. brucei equiperdum												[258]
081_EF1a_Trypanosoma_avium	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	AEIE---SKI	DRR---SGKE	L---EKNPK	AIKSGDAAL-	VRMVPQKPMC	VEVFNDYAPL	GRFAVRD---	[398]
082_EF1a_Trypanosoma_cruzi	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	AEIE---SKI	DRR---SGKE	L---EKNPK	SIKSGDAAM-	VRMVPQKPMC	VEVFNDYAPL	GRFAVRDMRQ	[419]
083_EF1a_Trypanosoma_rangeli	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	AEIE---SKI	DRR---SGKE	L---EKNPK	SIKSGDAAL-	VRMVPQKPMC	VEVFNDYAPL	GRFAVRDMRQ	[419]
084_EF1a_Ploeotia_costatum	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---NKI	DRR---SGKE	L---EAEPK	SIKSGDAAI-	VLMKPQKPMC	VESFTDYAPL	G-----	[409]
085_EF1a_Ploeotia_oblonga_cDNA	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---NKI	DRR---SGKE	L---EAEPK	SIKSGDAAI-	VLMKPQKPMC	VESFTDYAPL	G-----	[409]
086_EF1a_Ploeotia_oblonga_gDNA	IILN-HPGQI	SN-----	--GYAPVLD	CHTSHIACRF	AEIK---TKI	DRR---SGKE	L---EVEPK	AIKSGDAAI-	V-----			[389]
087_EF1a_Entosiphon_sp.	IVLN-HPGQI	SN-----	--GYAPVLD	CHTSHIACRF	AEIL---TKI	DRR---SGKE	L---EQEPK	AIKSGDAAM-	VVMRPLKPMC	VESFTDYAPL	G-----	[409]
088_EF1a_Entosiphon_sulcatum												[305]
089_EF1a_Peranema_tricophorum												[312]
090_EF1a_P. tricophorum FJ807243	IVLN-HPGQI	QN-----	--GYAPVLD	CHTSHIACRF	YEIK---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPTKPMC	VESFTDYAPL	GRFAVRDMRQ	[416]
091_EF1a_Colacium_elongatum	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYAPL	GRFAVRDMRQ	[418]
092_EF1a_Colacium_mucronatum	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGD---				[367]
093_EF1a_Colacium_sp.	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-		[407]
094_EF1a_Colacium_vesiculosum	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GR-----	[396]
095_EF1a_Cryptoglena_pigra	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[409]
096_EF1a_Cryptoglena_skujai	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[407]
097_EF1a_Discoplastis_adunca	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[409]
098_EF1a_Euglena_clara	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[378]
099_EF1a_Euglena_deses	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[409]
100_EF1a_Euglena_exilis	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[378]
101_EF1a_Euglena_geniculata	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[409]
102_EF1a_Euglena_gracilis	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIV---NKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDY??F	G-----	[409]
103_EF1a_E. gracilis_X16890	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GVSC-GDMRQ	[418]
104_EF1a_E. gracilis_FJ807238	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIV---NKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[419]
105_EF1a_Euglena_granulata	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[409]
106_EF1a_Euglena_hiemalis	IVLN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[409]
107_EF1a_Euglena_longa	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[257]
108_EF1a_Euglena_stellata	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[378]
109_EF1a_Euglena_viridis	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTD----		[390]
110_EF1a_Eutreptia_pertyi	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[409]
111_EF1a_Eutreptiella_pomquetensis	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---NKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFNDYPPF	G-----	[279]
112_EF1a_Lepocinlis_acus_var._major	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ANIQ---TKI	DRR---SGKE	L---EAEPK	HIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-		[407]
113_EF1a_Lepocinlis_ovum	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[378]
114_EF1a_Lepocinlis_spyrogyroides_clone_38	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	D-----	[180]
115_EF1a_L. spyrogyroides_clones_1_&_4	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	G-----	[409]
116_EF1a_Lepocinlis_steinii	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---NKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[378]
117_EF1a_Lepocinlis_tripteris	IVLN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPPF	G-----	[409]
118_EF1a_Monomorphina_aenigmatica	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYAPL	G-----	[409]
119_EF1a_Monomorphina_pyrum	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFNDYPP-	G-----	[409]
120_EF1a_Monomorphina_sp.	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESF?DY???	G-----	[409]
121_EF1a_Phacus_orbicularis	IVLN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	AEIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPC-			[382]
122_EF1a_Phacus_similis	IILN-HPGQI	SN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VES?TDYPP-		[407]
123_EF1a_Strombomonas_acuminata	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[419]
124_EF1a_Strombomonas_borysteniensis	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[378]
125_EF1a_Strombomonas_napiformis	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFNDYPPF	G-----	[409]
126_EF1a_Strombomonas_ovalis	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[418]
127_EF1a_Strombomonas_triquetra												[297]
128_EF1a_Trachelomonas_conradi	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPM-			[397]
129_EF1a_Trachelomonas_echinata	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-		[407]
130_EF1a_Trachelomonas_ilefvelei	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-		[407]
131_EF1a_Trachelomonas_plancunicica	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFAVRDMRQ	[378]
132_EF1a_Trachelomonas_similis	IILN-HPGQI	GN-----	--GYAPVLD	CHTSHIACRF	ATIQ---TKI	DRR---SGKE	L---EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	GRFA?RDMRQ	[378]

	450	460	470	480	490	500	510	520	530	540	550	
133_EFL1a_Trachelomonas_volvocina	IILN-HPGQI	GN-----	---GYAPVLD	CHTCHIACKF	ATIQ---TKI	DRR---SGKE	L----EAEPK	FIKSGDAAI-	VLMKPQKPMC	VESFTDYPP-	-----	[407]
134_EFL environmental_sample_EU810232	KVLT-HPNEL	KC-----	---GYCPIAF	IRTARSAVKM	VKIDWKIGKE	TGG---KK-	----LDDPV	SIKSNEMAQ-	LVFEPQPPFV	VDSFKSCEGL	GRVAIMEGNG	[449]
135_EFL Creolimax_fragrantissima												[331]
136_EFL Sphaeroforma_arctica	QVMN-HPGVL	KA-----	---GYCPIAF	VRTGRSAVKM	TSIDWKVGKE	TGG---TK-	----AENPA	GLKAADV-				[418]
137_EFL Monosiga_brevicollis	QIMN-HPGEL	KV-----	---GYCPIAF	VRTSRSAACRM	TAINWKIGKE	TGG---KK-	----AEDPV	ALKSNEVAE-	VVMAPQQPFV	VDTFKTCEGL	GRVAIMEGNS	[449]
138_EFL Xenopus_laevis	QIMN-HPGEL	KV-----	---GYCPVAF	VRTGRSAVKL	TKIAWKVGKE	TGG---SK-	----MENPQ	SIKANEMAE-	CTFEPMPQFV	VDSFKNCEGL	GRVALMEGNS	[449]
139_EFL Spizellomyces_punctatus	QTLN-IPGEV	KP-----	---GYSPIGF	VRCGRSACRM	KEIKWKMGKE	TGM---KK-	----LEGAH	SLKANEMAE-	VVFEPQIPLV	VDSFKNCEGL	SRIAFLDGN	[451]
140_EFL Gaertneriomyces_semiglobiferous												[168]
141_EFL Allomyces_macrogynus	QTLN-IPSEV	KP-----	---GYSPIGF	VRCGRSACKI	TKINWKIGKE	TGR---KK-	----MEGPH	SLKANEMAE-	VVFEPQIPLV	VDSFKNCEGL	SRIAFLDGN	[451]
142_EFL Blastocladiella_emersonii	QTLN-IPSEV	KP-----	---GYSPIGF	VRCGRSACKI	TKINWKIGKE	TGR---KK-	----MESPH	SLKANEMAE-	VVFEPQIPLV	VDTFGNCEGL	SRIAFLDGN	[451]
143_EFL Catenophylctis_sp.												[168]
144_EFL Coelomomyces_stegomyiae												[168]
145_EFL Entomophthora_muscae												[167]
146_EFL Basidiobolus_ranarum												[167]
147_EFL Conidiobolus_coronatus												[167]
148_EFL Tetrastelmis_tetrasthele	QTLN-IPGEI	KV-----	---GYSPIGF	VRCGRAACRI	MKLNVNKGKE	TGG---KM-	----LEDPH	SLKSNEVST-	IVFKPCQPLV	VDSYKNCEGL	SRIAFLDGN	[432]
149_EFL Mesostigma_viride												[293]
150_EFL Ostreococcus_tauri	QTLN-IPGEL	KV-----	---GYSPIGF	VRCGRSACKL	TALNFKVGKE	TGG---KK-	----LESPM	SLKSNEVAE-	CVFEPVHPLV	VDSFKKCEGL	SRIAFLDGN	[452]
151_EFL Nephrolepisma_olivacea												[319]
152_EFL Micromonas_pusilla	QVLD-VPQQL	KC-----	---GYSPVAF	VRCARSACKM	TELNVFKVGKE	TGG---KK-	----LESPM	ALKSNEVAE-	AVFEPQHPFV	VDHFKACDGL	ARIAFLDGN	[435]
153_EFL Parachlorella_kessleri												[314]
154_EFL Helicosporidium_sp.	QTLN-IPGEV	KA-----	---GYSPIGF	VRCGRAACRM	TKINWKVGKE	TGG---KK-	----LEDPH	SLKANEMAE-	VVFEPQIPLV	VDSFKNCEGL	SRIAFLDGN	[456]
155_EFL Gonium_pectorale	QTLN-IPGEV	KK-----	---GYSPIGF	VRCGRSACRM	QAINWKVGKE	TGG---KK-	----MEGPV	GLKANEMAE-	VVFEPQIPLV	VESAKSCEGL	SRIAFLDGN	[450]
156_EFL Chlamydomonas_reinhardtii	QTLN-IPGEV	KK-----	---GYSPIGF	VRCGRSACRI	SGINWKVGKE	TGG---KK-	----MENPV	GLKANEMAE-	VVFEPQIPLV	VDSFKNCEGL	SRIAFLDGN	[450]
157_EFL Scenedesmus_obliquus	QTLN-IPGEV	KA-----	---GYSPIGF	VRCGRSACKI	SKIDWKVGKE	TGG---KK-	----LEAPH	ALKANEMAE-	VVFEPQIPLV	VDSFKSCEGL	SRIAFLDGN	[450]
158_EFL "Oryza_sativa"	QTLN-IPGEV	KK-----	---GYSPIGF	VRCGRAACKI	TEIKWKVGKE	TGN---KK-	----LEEPP	GLKANEMAE-	VVFEPQIPLV	VDSFKSCEGL	SRIAFLDGN	[450]
159_EFL Chlorococcum_sp.	QTLN-IPGEL	KV-----	---GYSPIGF	VRCGRAACRV	TNLDWKVGKE	TGG---KK-	----LEEPP	SLKSNEMAQ-	IVFTPCQPLV	VDSFKNCEGL	SRIAFLDGN	[433]
160_EFL Oclochaete_hystrix	QTLN-IPGEV	KA-----	---GYSPIGF	VRCGRSACRI	SGINWKVGKE	TGG---KK-	----LEAPH	SLKANEMAE-	VVFEPQIPLV	VDSFKNCEGL	SRIAFLDGN	[423]
161_EFL Ulva_intestinalis	QTLN-IPGEL	KA-----	---GYSPIGF	VRCGRSACRV	TKIDWKVGKE	TGG---KK-	----LEEPP	SLKSNEMAQ-	IVFKPAQPLV	VDTFKNCEGL	SRIAFLDGN	[433]
162_EFL Urospora_sp.	QTLN-IPGEL	KV-----	---GYSPIGF	VRCGRSACRV	SAIDWKVGKE	TGG---KK-	----LDEPH	SLKSNEMAQ-	VVFEPQIPLV	VDSFKNCEGL	SRIAFLDGN	[362]
163_EFL Gymnochlorella_stellata_ppc	QVLD-VPHEV	KP-----	---GYSPIGF	VRCGRSACKL	KKINWKMGKE	TGG---KK-	----VENPA	MLKHGEMAE-	VVFEPMFGEV	AEPFKRCEGL	ARIAFLDGN	[462]
164_EFL Lotharella_amoebiformis_ppc	QVLD-VPHEI	KP-----	---GYSPIGF	VRCGRSACRV	KKINWKMGKE	TGG---KK-	----LEDPK	SLKHGEMAE-	VVFEPMFGEV	TESFKTCEGL	ARIAFLDGN	[377]
165_EFL Chlorarachnion_reptans_ppc	QVLD-VPHEI	KP-----	---GYSPIGF	VRCGRSACRV	KKINWKMGKE	TGG---KK-	----LENPP	HLKHGEMAE-	VVFEPMFGEV	AEGFKSCEGL	ARIAFLDGN	[386]
166_EFL Chondrus_crispus	QTLN-HPLYF	KI-----	---GYAPVAF	VRTGRSACRI	KKINWKMGKE	TGG---AK-	----VENPD	YIKANEVAE-	IVFAPQPPFV	VEGFKKCEGL	GRVAIM----	[445]
167_EFL Gracilaria_changii	QTLN-HPGQL	KV-----	---GYAPVAF	VRTGRSACKI	KSINWKMGKE	TGG---AK-	----VEHPE	YIKANEVAE-	IVFEPQPPFV	VEGFKKCEGL	GRVAIMEGNT	[449]
168_EFL Cryptomonas_ovata	QTLNIPGEL	KV-----	---GYSPIGF	VRCGRAACRM	SVIDWKMGKE	TGG---AK-	----AESPP	HLKANEVAQ-	AQFVPMTPLV	CDTFKNCEGL	SRIAFLDGN	[435]
169_EFL Guillardia_theta	QTLNIPGEL	KT-----	---GYSPIGF	VRCGRAACRM	TVIDWKMGKE	TGG---QK-	----LENPP	HLKANEVAQ-	AQFEPMTPLV	CDTFKNCEGL	SRIAFLDGN	[451]
170_EFL Rhodomonas_salina	QTLNIPGEL	KV-----	---GYSPIGF	VRCGRAACRM	SVIDWKMGKE	TGG---QK-	----LESPH	HLKANEVAQ-	AQFEPMTPLV	CDTFKNCEGL	SRIAFLDGN	[450]
171_EFL Pavlova_lutheri	QTLN-IPGEL	KV-----	---GYSPIGF	VRCGRAACRI	TEIKFKVGKE	TG-----						[402]
172_EFL Chrysochromulina_sp.	QTLN-IPGEI	KL-----	---GYSPIGF	VRCGRSACRA	TGFNFKVGK-							[339]
173_EFL Isochrysis_galbana	QTLN-IPGEI	KI-----	---GYSPIGF	VRCGRAACRA	TAFVFKVGKE	TGG---KK-	----MEAPH	SLKSNEMAQ-	VSRFPQIPLV	VDSFKNCEGL	SRIAFLDGN	[450]
174_EFL Emiliana_huxleyi	QTLN-IPGEI	KL-----	---GYSPIGF	VRCGRSACRA	TAFKFKVGKE	TGG---KK-	----MESPH	SLKSNEMAQ-	VSRFPQIPLV	VDSFKNCEGL	SRIAFLDGN	[452]
175_EFL Achnanthes_kuwaitensis	AVQE-HPGVL	KP-----	---GYCPVVF	CRTAKVACKM	TKILWKQSKK	TGG---AK-	----VENPP	ELSQYENAE-	VEFEPTAPLF	VEPFETCAAL	GRIAVMDSNR	[419]
176_EFL Asterionella_glacialis	AVQE-HPGVL	KP-----	---GYCPVVF	CRTAKVACKM	TKILWKQSKK	TGG---EK-	----VENPP	ELSQYENAE-	VEFEPTAPLF	LEPFETCAAL	GRIAVMDSNR	[419]
177_EFL Detonula_confervacea	AVQE-HPGVL	KA-----	---GYAPVVF	CRTAKVACKM	SKILWKMGKK	TGG---QK-	----LDNPT	ELSQYENAE-	VEFVPTAPLF	VEPFETCAAL	GRIAVMDSNR	[419]
178_EFL Thalassiosira_pseudonana	AVQE-HPGVL	KP-----	---GYAPVVF	CRTAKVACKM	TKILWKMGKK	TGG---QK-	----VDNPP	ELSQYENAE-	VEFEPSQPLV	VEPFETCAAL	GRIAVMDSNR	[440]
179_EFL Perkinsus_marinus	QVLD-VPGEI	KV-----	---GYSPIAF	VRTGRSACRL	AGIKWKVGKE	TGG---KK-	----MEDPH	ALKSNEMAE-	VVFEPQIPLV	VDSFKNCEGL	SRIAFLDGN	[455]
180_EFL Oxyrrhis_marina	QVLD-IPGEI	KM-----	---GYSPIGF	VRCGRAACRI	SGLNWKVGKE	TGG---KK-	----MPEPH	SLKSNEMAE-	VVFSPMQPLV	VDTFKNCEGL	SRIAFLDGN	[453]
181_EFL Karlodinium_micrum	QVLD-IPNEI	KV-----	---GYSPIGF	VRCGRSACRI	SKLWKMGKE	TGG---KK-	----MEDPH	ALKSNEMAE-	ASFQPPQPLV	CDTFKNCEGL	SRIAFLDGN	[451]
182_EFL Amphidinium_carterae	QVLD-IPNEI	KV-----	---GYSPIGF	VRCGRAACRI	SALKWKMGKE	TGG---KK-	----MEDPH	SLKSNEMAE-	CTFQPPQPLV	CDTFKNCEGL	SRIAFLDGN	[450]
183_EFL Heterocapsa_triquetra	QVLD-IPNEI	KT-----	---GYSPIGF	VRCGRAACRI	SALKWKMGKE	TGG---KK-	----MEDPH	SLKSNEMAE-	CAFQPPQPLV	CDTFKNCEGL	SRIAFLDGN	[451]
184_EFL Planoglabratella_operculis	FVQD-HPGKL	A-----										[319]
185_EFL Reticulomyxa_filosa	FVQD-HPGQL	KAATNDGKGN	YKGGFTPSIH	IRTAKAPQM	LEIKWKMGKA	TNN---AK-	----VEGAT	FIEAGDQAE-	VVFAPKMPLV	VTSFDECKPL	GRMAAMDSNS	[431]
186_EFL Lotharella_amoebiformis	KVQE-HPGQL	KV-----	---GYCPVHL	VRTAKSACKL	SKIVWKVTKA	NLKQCKKKT	MENYKEMNP	FIKAGDQCE-	IEFEPQGPLS	VSPFKCEGL	GRIAVLESNS	[486]
187_EFL Gymnochlorella_stellata	KIQE-HPGKL	KV-----	---GYTPLVL	IRTAKAPCRV	SKILWKVTKA	NLTKIKKAE	MENFKEKAD	FVQKGDVAE-	IEFEPQMPMS	VSTLQKCEGL	GRVALESNS	[477]
188_EFL Chlorarachnion_reptans	KIQE-HPGKL	KV-----	---GYTPLVL	IRTAKAPCRV	SKILWKVTKA	NLTKIKKAE	MENFKEKAD	FVQKGDVAE-	IEFEPQMPIS	VSTLQKCEGL	GRVALESNS	[454]
189_EFL Diplonema_ambulator	QVLD-HPGEL	KV-----	---GYTPVCF	VRTGRSAVRM	MELVWKMGKE	TGG---KK-	----AEAPH	CLKANEMAE-	VVFAPQIPLV	VDSFKNCEGL	SRIAFLDGN	[317]
190_EFL Diplonema_papillatum	QVLD-HPGEL	KV-----	---GYTPVCF	VRTGRSAVRM	SQLVWKMGKE	TGG---KK-	----AEAPH	CLKANEMAE-	VVFAPQIPLV	VDTFKNCEGL	SRIAFLDGN	[441]
191_EFL Rhynchopus_euleeides	QVLD-HPGEL	KV-----	---GYTPVCF	VRTGRSAVRM	TQLVWKMGKE	TGG---KK-	----AEAPH	CLKANEMAE-	VVFAPQIPLV	VDTFKNCEGL	SRIAFLDGN	[322]
192_EFL Petalomonas_cantuscygni												[308]
193_EFL Trypanoplasma_borrel												[339]
194_EFL Neobodo_salien	MVLD-HPGQL	RV-----	---GYTPCTY	VRTSHVACRL	SEIRFKAGKK	TGG---QK-	----VESPE	FIEKGMCE-	LVWKPQPPFV	VDDFKS----		[426]

	560	570	580	590	600	
001_aEF1a_Desulfurococcus_mobilis	TIGIGQVLEI	KPAQVNIKK-	-----	-----	-----	[438]
002_aEF1a_Sulfolobus_solfataricus	TVGVGIIVDV	KPAKVEIK--	-----	-----	-----	[435]
003_aEF1a_Thermoplasma_acidophilum	TVAAGQCIDL	EKR-----	-----	-----	-----	[427]
004_aEF1a_Halobacterium_sp.	TIAAGQVLDV	DEA-----	-----	-----	-----	[421]
005_aEF1a_Methanocaldococcus_jannaschii	TIAAGMAIDV	KAKNK-----	-----	-----	-----	[428]
006_EF1a_Acanthamoeba_culbertsoni	TVAVGIL---	-----	-----	-----	-----	[415]
007_EF1a_Neoparamoeba_branchipila	-----	-----	-----	-----	-----	[400]
008_EF1a_Planoprotostelium_aurantium	-----	-----	-----	-----	-----	[401]
009_EF1a_Physarum_polycephalum	-----	-----	-----	-----	-----	[401]
010_EF1a_Amoebidium_parasiticum	-----	-----	-----	-----	-----	[289]
011_EF1a_Capsaspora_owczaraki	TVAVGVIKAV	DKSAKDAKAD	GKDKKAAAPA	KAAKK-----	-----	[464]
012_EF1a_Monosiga_ovata	TVAVGI---	-----	-----	-----	-----	[428]
013_EF1a_Salpingoeca_amphoridium	TVAVGIL---	-----	-----	-----	-----	[430]
014_EF1a_Ceanorhabditis_elegans	TVAVGVIKSV	EKSDGSSGKV	TKSAQKAAPK	KK-----	-----	[463]
015_EF1a_Drosophila_melanogaster	TVAVGVIKAV	NFKDASGGKV	TKAAEKATKG	KK-----	-----	[463]
016_EF1a_Ornithorhynchus_anatinus	TVAVGVIKNV	EKSSGGAGKV	TKSAQKAQKA	GK-----	-----	[463]
017_EF1a_Xenopus_laevis	TVAVGVIKGV	DKKLASSGKV	TKSAAKAGKK	-----	-----	[461]
018_EF1a_Saccharomyces_cerevisiae	TVAVGVIKSV	DKTEKAAKVT	KAAQKAAKK-	-----	-----	[458]
019_EF1a_Neurospora_crassa	TVAVGVIKAV	DKSTAAAGKV	TKSAAKAAKK	-----	-----	[460]
020_EF1a_Cryptococcus_neoformans	TVAVGVIKSV	DKTEKGGKVT	KAAEKAAKKK	-----	-----	[459]
021_EF1a_Schizophyllum_commune	TVAVGVIKSV	DKTDKSGGKV	TKSAEKAALK	K-----	-----	[460]
022_EF1a_Monoblepharis_insignis	-----	-----	-----	-----	-----	[390]
023_EF1a_Rhizophlyctis_harderi	-----	-----	-----	-----	-----	[390]
024_EF1a_Harpochytrium_sp.	-----	-----	-----	-----	-----	[359]
025_EF1a_Absidia_glauca	TVAVGVIKAV	EKVDKAGKVT	KAAAKAGEK-	-----	-----	[458]
026_EF1a_Smittium_simulii	TVAVGIL---	-----	-----	-----	-----	[427]
027_EF1a_Chara_australis	TVAVGVIKAV	EKKEKEGKVT	KAAKKGGKK-	-----	-----	[431]
028_EF1a_Physcomitrella_patens	TVAVGVIKAV	EKKEPTGAKV	TKAAAKKK--	-----	-----	[447]
029_EF1a_Acetabularia_acetabulum	TVAVGVIKKV	EKKDPSAAKT	TKAAAKKK--	-----	-----	[430]
030_EF1a_Chaetomorpha_coliformis	TVAVGVIKEV	EKKEVAGKTT	KAAAKKK--	-----	-----	[379]
031_EF1a_Arabidopsis_thaliana	TVAVGVIKSV	DKKDPTGAKV	TKAAVKKGAK	-----	-----	[449]
032_EF1a_Oryza_sativa	TVAVGVIKNV	EKKDPTGAKV	TKAAAKKK--	-----	-----	[447]
033_EF1a_Cyanidiodioschyzon_merolae	TVAVGVIKEV	NKKEAEGKVT	KAAAKKA---	-----	-----	[450]
034_EF1a_Porphyrha_yezoensis	TVAVGVIKSV	EKKEVEGKMT	KSAKK-----	-----	-----	[449]
035_EF1a_Plasmodium_falciparum	TIAGGIKSV	EKKEPGAVTA	KAPAKK----	-----	-----	[443]
036_EF1a_Cryptosporidium_parvum	TVAVGVIKSV	TKKEATSKKK	-----	-----	-----	[435]
037_EF1a_Tetrahymena_pyriformis	TVAVGVIKKV	EKKDK-----	-----	-----	-----	[435]
038_EF1a_Stentor_coeruleus	TVAV-----	-----	-----	-----	-----	[409]
039_EF1a_Phytophthora_parasitica	TVAVGVIKSV	TKKDASAKGA	AKKK-----	-----	-----	[443]
040_EF1a_Mallomonas_rasilis	TVAVGVIK--	-----	-----	-----	-----	[414]
041_EF1a_Phaeodactylum_tricornutum	TVAVGVIKAT	TKEEGGKKKK	-----	-----	-----	[439]
042_EF1a_Thalassiosira_pseudonana	TVAVGVIKAV	E-----	-----	-----	-----	[430]
043_EF1a_Malawimonas_jakobiformis	-----	-----	-----	-----	-----	[395]
044_EF1a_Malawimonas_californiana	TVAVGVIKAV	TKTEGGKKAG	AAATKKK---	-----	-----	[446]
045_EF1a_Trimastix_pyriformis	-----	-----	-----	-----	-----	[404]
046_EF1a_Dinenympha_exilis	-----	-----	-----	-----	-----	[395]
047_EF1a_Monocercomonoides_sp.	TVAVGVIRSV	NKKSRAKVT	KAAAKAQKK-	-----	-----	[448]
048_EF1a_Strebloamastix_strix	-----	-----	-----	-----	-----	[398]
049_EF1a_Histomonas_meleagridis	TVAVGVIRAV	EKKVVEDKKK	K-----	-----	-----	[440]
050_EF1a_Trichomonas_vaginalis	TVAVGVIRSV	NKKPNPIK--	-----	-----	-----	[437]
051_EF1a_Hexamita_inflata	TVAVGVVTEV	LKKDK-----	-----	-----	-----	[418]
052_EF1a_Spironucleus_vortens	-----	-----	-----	-----	-----	[400]
053_EF1a_Andalucia_incarcerata	-----	-----	-----	-----	-----	[401]
054_EF1a_Seculamonas_ecuadoriensis	TVAVGVIKSV	EKKETGAGKA	GKAAGKKK--	-----	-----	[447]
055_EF1a_Reclinomonas_americana	TVAVGVIKEV	DKMAEKSGKA	AAGKAGGAAG	KGK-----	-----	[452]
056_EF1a_Histiona_roides	TVAVGVIKEV	AKVADKAAGG	KAGKKK----	-----	-----	[445]
057_EF1a_Jakoba_libera	TVAVGVIKSV	EKREYSPKVG	GKKK-----	-----	-----	[443]
058_EF1a_Stachyamoeba_lipophora	TVAVGIKSV	TKKAK-----	-----	-----	-----	[433]
059_EF1a_Naegleria_gruberi	-----	-----	-----	-----	-----	[402]
060_EF1a_Naegleria_andersoni	-----	-----	-----	-----	-----	[405]
061_EF1a_Acrasis_rosea	-----	-----	-----	-----	-----	[401]
062_EF1a_Sawyeria_marylandensis	TVAVGVIKKV	EKKDKF----	-----	-----	-----	[435]
063_EF1a_Bodo_saltans	-----	-----	-----	-----	-----	[369]
064_EF1a_Rhynchobodo_sp.	-----	-----	-----	-----	-----	[367]
065_EF1a_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	[404]
066_EF1a_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	[375]

	560	570	580	590	600	
067_EF1a_Perkinsiella_amoebae	-----	-----	-----	-----	-----	[305]
068_EF1a_Leishmania_major	TVAVGIIKGV	NKKEGSGGKV	TKAAAKASKK	-----	-----	[449]
069_EF1a_Leishmania_infantum	TVAVGIIKGV	NKKEGSGGKV	TKAAAAAASKK	-----	-----	[449]
070_EF1a_Leishmania_braziliensis	TVAVGIIKAV	SKKDGSAGKV	TKAAAAAASKK	-----	-----	[449]
071_EF1a_Leishmania_tarentolae	-----	-----	-----	-----	-----	[344]
072_EF1a "L. tarentolae" FJ807257	-----	-----	-----	-----	-----	[305]
073_EF1a_Leptomonas_costaricensis	-----	-----	-----	-----	-----	[198]
074_EF1a_Leptomonas_bifurcata	-----	-----	-----	-----	-----	[277]
075_EF1a_Leptomonas_podlipaevi	-----	-----	-----	-----	-----	[271]
076_EF1a_Herpetomonas_muscarum	-----	-----	-----	-----	-----	[381]
077_EF1a_Herpetomonas_pessoai	-----	-----	-----	-----	-----	[257]
078_EF1a_Blastocrithidia_culicis	-----	-----	-----	-----	-----	[255]
079_EF1a_Trypanosoma_brucei	TVAVGIIKAV	TKKDGSAGKV	TKAAVKASKK	-----	-----	[449]
080_EF1a T. brucei equiperdum	-----	-----	-----	-----	-----	[258]
081_EF1a_Trypanosoma_avium	-----	-----	-----	-----	-----	[398]
082_EF1a_Trypanosoma_cruzi	TVAVGIIKAV	TKKEGGAGKV	TKAAAKASKK	-----	-----	[449]
083_EF1a_Trypanosoma_rangeli	TVAVGIIKAV	TKKEGGAGKV	TKAAAAAASKK	-----	-----	[449]
084_EF1a_Ploeotia_costatum	-----	-----	-----	-----	-----	[409]
085_EF1a_Ploeotia_oblonga_cDNA	-----	-----	-----	-----	-----	[409]
086_EF1a_Ploeotia_oblonga_gDNA	-----	-----	-----	-----	-----	[389]
087_EF1a_Entosiphon_sp.	-----	-----	-----	-----	-----	[409]
088_EF1a_Entosiphon_sulcatum	-----	-----	-----	-----	-----	[305]
089_EF1a_Peranema_tricophorum	-----	-----	-----	-----	-----	[312]
090_EF1a P. tricophorum FJ807243	TVAVGVIIKAV	NKKETSGKVT	KAAQKKK	---	-----	[443]
091_EF1a_Colacium_elongatum	HV-----	-----	-----	-----	-----	[420]
092_EF1a_Colacium_mucronatum	-----	-----	-----	-----	-----	[367]
093_EF1a_Colacium_sp.	-----	-----	-----	-----	-----	[407]
094_EF1a_Colacium_vesiculosum	-----	-----	-----	-----	-----	[396]
095_EF1a_Cryptoglena_pigra	-----	-----	-----	-----	-----	[409]
096_EF1a_Cryptoglena_skujai	-----	-----	-----	-----	-----	[407]
097_EF1a_Discoplastis_adunca	-----	-----	-----	-----	-----	[409]
098_EF1a_Euglena_clara	TVAVGIL---	-----	-----	-----	-----	[385]
099_EF1a_Euglena_deses	TVAVGIL---	-----	-----	-----	-----	[409]
100_EF1a_Euglena_exilis	TVAVGIL---	-----	-----	-----	-----	[385]
101_EF1a_Euglena_geniculata	-----	-----	-----	-----	-----	[409]
102_EF1a_Euglena_gracilis	-----	-----	-----	-----	-----	[409]
103_EF1a E. gracilis X16890	TVAVGVIIKSV	NKKENTGKVT	KAAQKKK	---	-----	[445]
104_EF1a E. gracilis FJ807238	TVAVGVIIKSV	NKKENTGKVT	KAAQKKK	---	-----	[446]
105_EF1a_Euglena_granulata	-----	-----	-----	-----	-----	[409]
106_EF1a_Euglena_hiemalis	-----	-----	-----	-----	-----	[409]
107_EF1a_Euglena_longa	TVAVGVIIKAV	NKKENTGKVT	KAAQKKK	---	-----	[284]
108_EF1a_Euglena_stellata	TVAVGIL---	-----	-----	-----	-----	[385]
109_EF1a_Euglena_viridis	-----	-----	-----	-----	-----	[390]
110_EF1a_Eutreptia_pertyi	-----	-----	-----	-----	-----	[409]
111_EF1a_Eutreptiella_pomquetensis	-----	-----	-----	-----	-----	[279]
112_EF1a_Lepocinlis_acus_var._major	-----	-----	-----	-----	-----	[407]
113_EF1a_Lepocinlis_ovum	TVAVGIL---	-----	-----	-----	-----	[385]
114_EF1a_Lepocinlis_spyrogyroides_clone_38	-----	-----	-----	-----	-----	[180]
115_EF1a L. spyrogyroides clones_1_&_4	-----	-----	-----	-----	-----	[409]
116_EF1a_Lepocinlis_steinii	TVAVGIL---	-----	-----	-----	-----	[385]
117_EF1a_Lepocinlis_tripteris	-----	-----	-----	-----	-----	[409]
118_EF1a_Monomorphina_aenigmatica	-----	-----	-----	-----	-----	[409]
119_EF1a_Monomorphina_pyrum	-----	-----	-----	-----	-----	[409]
120_EF1a_Monomorphina_sp.	-----	-----	-----	-----	-----	[409]
121_EF1a_Phacus_orbicularis	-----	-----	-----	-----	-----	[382]
122_EF1a_Phacus_similis	-----	-----	-----	-----	-----	[407]
123_EF1a_Strombomonas_acuminata	TVAVGVIIKAV	NKKEGSGGKV	KAAPKKK	---	-----	[446]
124_EF1a_Strombomonas_borysteniensis	TVAVGIL---	-----	-----	-----	-----	[385]
125_EF1a_Strombomonas_napiformis	-----	-----	-----	-----	-----	[409]
126_EF1a_Strombomonas_ovalis	TVAVGIL---	-----	-----	-----	-----	[425]
127_EF1a_Strombomonas_triquetra	-----	-----	-----	-----	-----	[297]
128_EF1a_Trachelomonas_conradi	-----	-----	-----	-----	-----	[397]
129_EF1a_Trachelomonas_echinata	-----	-----	-----	-----	-----	[407]
130_EF1a_Trachelomonas_lefevrei	-----	-----	-----	-----	-----	[407]
131_EF1a_Trachelomonas_planc-tonica	TVAVGIL---	-----	-----	-----	-----	[385]
132_EF1a_Trachelomonas_similis	TVAVGIL---	-----	-----	-----	-----	[385]

	560	570	580	590	600	
133_EFL1a_Trachelomonas_volvocina	-----	-----	-----	-----	-----	[407]
134_EFL environmental_sample_EU810232	VVMLGKVVSV	TFKE-----	-----	-----	-----	[463]
135_EFL Creolimax_fragrantissima	-----	-----	-----	-----	-----	[331]
136_EFL Sphaeroforma_arctica	-----	-----	-----	-----	-----	[418]
137_EFL Monosiga_brevicollis	VVMLGKVTKV	ERA-----	-----	-----	-----	[462]
138_EFL Xenopus_laevis	VVMLGKVVTV	EFL-----	-----	-----	-----	[462]
139_EFL Spizellomyces_punctatus	AVMLGKVTQV	EYKD-----	-----	-----	-----	[465]
140_EFL Gaertneriomyces_semiglobiferous	-----	-----	-----	-----	-----	[168]
141_EFL Allomyces_macrognus	AVMLGKVVAV	EHK-----	-----	-----	-----	[464]
142_EFL Blastocladia_emersonii	AVMLGKVVAV	EHK-----	-----	-----	-----	[464]
143_EFL Catenophylctis_sp.	-----	-----	-----	-----	-----	[168]
144_EFL Coelomomyces_stegomyiae	-----	-----	-----	-----	-----	[168]
145_EFL Entomophthora_muscae	-----	-----	-----	-----	-----	[167]
146_EFL Basidiobolus_ranarum	-----	-----	-----	-----	-----	[167]
147_EFL Conidiobolus_coronatus	-----	-----	-----	-----	-----	[167]
148_EFL Tetraselmis_tetrathele	AVMMGKVMAS	TMSA-----	-----	-----	-----	[446]
149_EFL Mesostigma_viride	-----	-----	-----	-----	-----	[293]
150_EFL Ostreococcus_tauri	AVMLGKITNV	EYKAI-----	-----	-----	-----	[467]
151_EFL Nephroselmis_olivacea	-----	-----	-----	-----	-----	[319]
152_EFL Micromonas_pusilla	AVMLGKITKV	QFKEDK-----	-----	-----	-----	[451]
153_EFL Parachlorella_kessleri	-----	-----	-----	-----	-----	[314]
154_EFL Helicosporidium_sp.	AVMLGKVST	TSK-----	-----	-----	-----	[469]
155_EFL Gonium_pectorale	AVMLGKVTDV	LFK-----	-----	-----	-----	[463]
156_EFL Chlamydomonas_reinhardtii	AVMLGKVVSV	THK-----	-----	-----	-----	[463]
157_EFL Scenedesmus_obliquus	AVMLGKVTV	TFK-----	-----	-----	-----	[463]
158_EFL "Oryza_sativa"	AVMLGKVAV	THK-----	-----	-----	-----	[463]
159_EFL Chlorococcum_sp.	AVMLGKVVKV	THK-----	-----	-----	-----	[446]
160_EFL Ochlochaete_hystrix	-----	-----	-----	-----	-----	[423]
161_EFL Ulva_intestinalis	AVMLGKVVKV	THK-----	-----	-----	-----	[446]
162_EFL Urospora_sp.	-----	-----	-----	-----	-----	[362]
163_EFL Gymnochlora_stellata_ppc	VVMLGKVTKT	VPKSFSGDK- KDKGGKDG- DKKGKDDKKK DDKKSSSKKK K	-----	-----	-----	[510]
164_EFL Lotharella_amoebiformis_ppc	VVMLGKVTKT	TPKQFADPK- KKAGSGDKKK DDKKDDKKK SAKKKK----	-----	-----	-----	[422]
165_EFL Chlorarachnion_reptans_ppc	VVMLGKVTKT	QPKSYGNKGD KKGDKKGDGK DDKKDDKKKS TKKKK-----	-----	-----	-----	[431]
166_EFL Chondrus_crispus	-----	-----	-----	-----	-----	[445]
167_EFL Gracilaria_changii	VVMLGKAIV	EFYT-----	-----	-----	-----	[463]
168_EFL Cryptomonas_ovata	-----	-----	-----	-----	-----	[435]
169_EFL Guillardia_theta	VVMLGKVIAT	VARDGGAKK K-----	-----	-----	-----	[472]
170_EFL Rhodomonas_salina	VVMLGKVIAT	VAREDAGKKK -----	-----	-----	-----	[470]
171_EFL Pavlova_lutheri	-----	-----	-----	-----	-----	[402]
172_EFL Chrysochromulina_sp.	-----	-----	-----	-----	-----	[339]
173_EFL Isochrysis_galbana	AVMLGKVVSV	IAKQDMPPEE DKKKKK-----	-----	-----	-----	[476]
174_EFL Emiliana_huxleyi	AVMLGKVVSV	IQKKDMPPEE GKKKK-----	-----	-----	-----	[477]
175_EFL Achnanthes_kuwaitensis	LKMLGKVTGV	VSSTA-----	-----	-----	-----	[434]
176_EFL Asterionella_glacialis	LKMLGKVTHT	ES-----	-----	-----	-----	[431]
177_EFL Detonula_confervacea	LKMLGKVTGT	TTVE-----	-----	-----	-----	[433]
178_EFL Thalassiosira_pseudonana	LKMLGKVTGV	TSV-----	-----	-----	-----	[453]
179_EFL Perkinsus_marinus	VVMLGKCVST	VAKDVK-----	-----	-----	-----	[471]
180_EFL Oxyrrhis_marina	AVMLGKIVS?	VDKGVDGKA GKAGKKK----	-----	-----	-----	[480]
181_EFL Karodinium_micrum	VVMLGKVISC	EKKDENDKGK GKKK-----	-----	-----	-----	[475]
182_EFL Amphidinium_carterae	-----	-----	-----	-----	-----	[450]
183_EFL Heterocapsa_triquetra	VVMLGKVISC	ERKEEGDKGK KKK-----	-----	-----	-----	[474]
184_EFL Planoglabratella_operculis	-----	-----	-----	-----	-----	[319]
185_EFL Reticulomyxa_filosa	LMLGKVVSV	EYK-----	-----	-----	-----	[444]
186_EFL Lotharella_amoebiformis	LVMLGKINSV	VNTEIK-----	-----	-----	-----	[502]
187_EFL Gymnochlora_stellata	LVMLGKIQSA	EHGAVKDLA-----	-----	-----	-----	[496]
188_EFL Chlorarachnion_reptans	LVMLGKIATA	EQTAVADRT-----	-----	-----	-----	[473]
189_EFL Diplonema_ambulator	VVMLGKVTKT	SEKVAAAKK-----	-----	-----	-----	[337]
190_EFL Diplonema_papillatum	VVMLGKVVKT	AEKAKDDGKG KKK-----	-----	-----	-----	[464]
191_EFL Rhynchopus_euleeides	VVMLGKVTKT	TEKVAGDKKG KK-----	-----	-----	-----	[344]
192_EFL Petalomonas_cantuscygni	-----	-----	-----	-----	-----	[308]
193_EFL Trypanoplasma_borreli	-----	-----	-----	-----	-----	[339]
194_EFL Neobodo_salien	-----	-----	-----	-----	-----	[426]

Figure 1.4. The cladogram summary of the RAxML ML nucleotide gene tree with clade recovery shown using the MolcGo3rd alignment with 194 OTUs. Recovered relationships are denoted with ++ at the nodes if all four analysis types (RAxML, TNT, GARLI, MrBayes) recover them or a + if only three of the four types recover them. Any relationships recovered by two or less analysis types are not labeled. Gene recovery is denoted the same way, but at the end of the gene name instead of at a node. Colors are indicative of supergroup of the organism in the tree: no color=Archaea, purple=Amoebozoa, green=Archaeplastida, orange=Chromalveolata, blue=Excavata, yellow=Opisthokonta, and pink=Rhizaria. Unikonts are a combination of the supergroups Amoebozoa and Opisthokonta and this clade is represented by alternating colors of those supergroups (purple and yellow) in the EF1 α portion of the gene tree. Genes that are recovered as monophyletic are denoted with vertical bars with the gene name to the right.

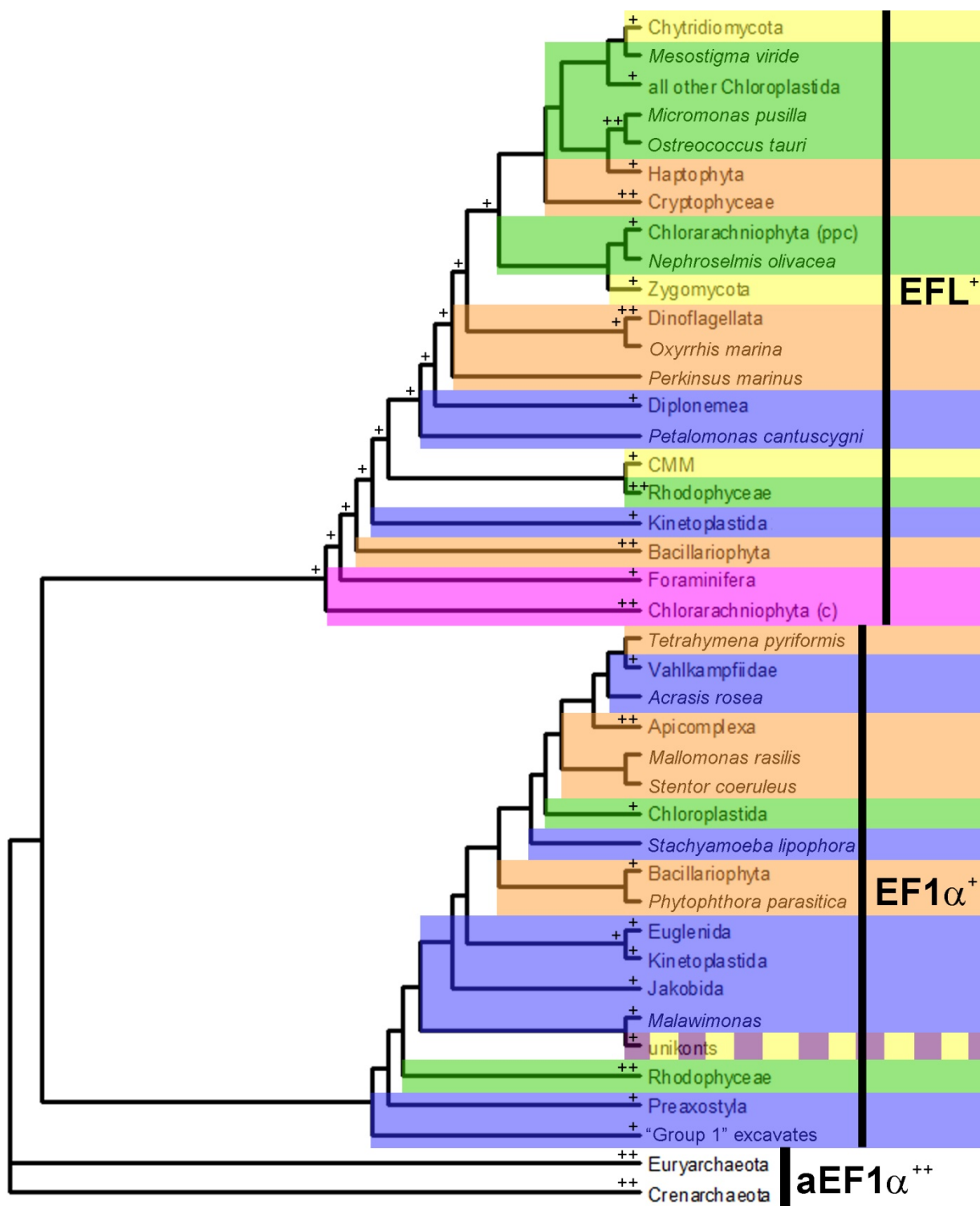


Figure 1.5. The detailed EF1 α portion of the nucleotide gene tree with its representative OTUs. This is the RAxML ML tree with bootstrap support using the MolcGo3rd alignment with 194 OTUs with the EFL and aEF1 α portions of the gene tree summarized. The value on the top is the RAxML bootstrap score and the values below are the GARLI bootstrap score before the hash mark and the MrBayes posterior probability after. Only bootstraps $\geq 50\%$ and posterior probabilities ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bootstrap support from RAxML and GARLI as well as 1.00 posterior probability from the MrBayes resulting tree topologies for the same alignment. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the support thresholds stated. Colors are indicative of supergroup of the organism in the tree: no color=Archaea, purple=Amoebozoa, green=Archaeplastida, orange=Chromalveolata, blue=Excavata and yellow=Opisthokonta. Groups that are recovered as monophyletic are denoted with vertical bars with the clade name to the right with relatively deep level clades to the right and relatively tip level clades closer to the OTUs. The star marks the OTU that still retains the endogenous EF1 α and EFL. **A.** The Euglenozoa portion of the EF1 α tree with a closed black circle indicating where it connects to the rest of the EF1 α detailed tree. **B.** The rest of the eukaryotic EF1 α detailed tree with the closed black circle indicating where it connects to the Euglenozoa OTUs.

B.

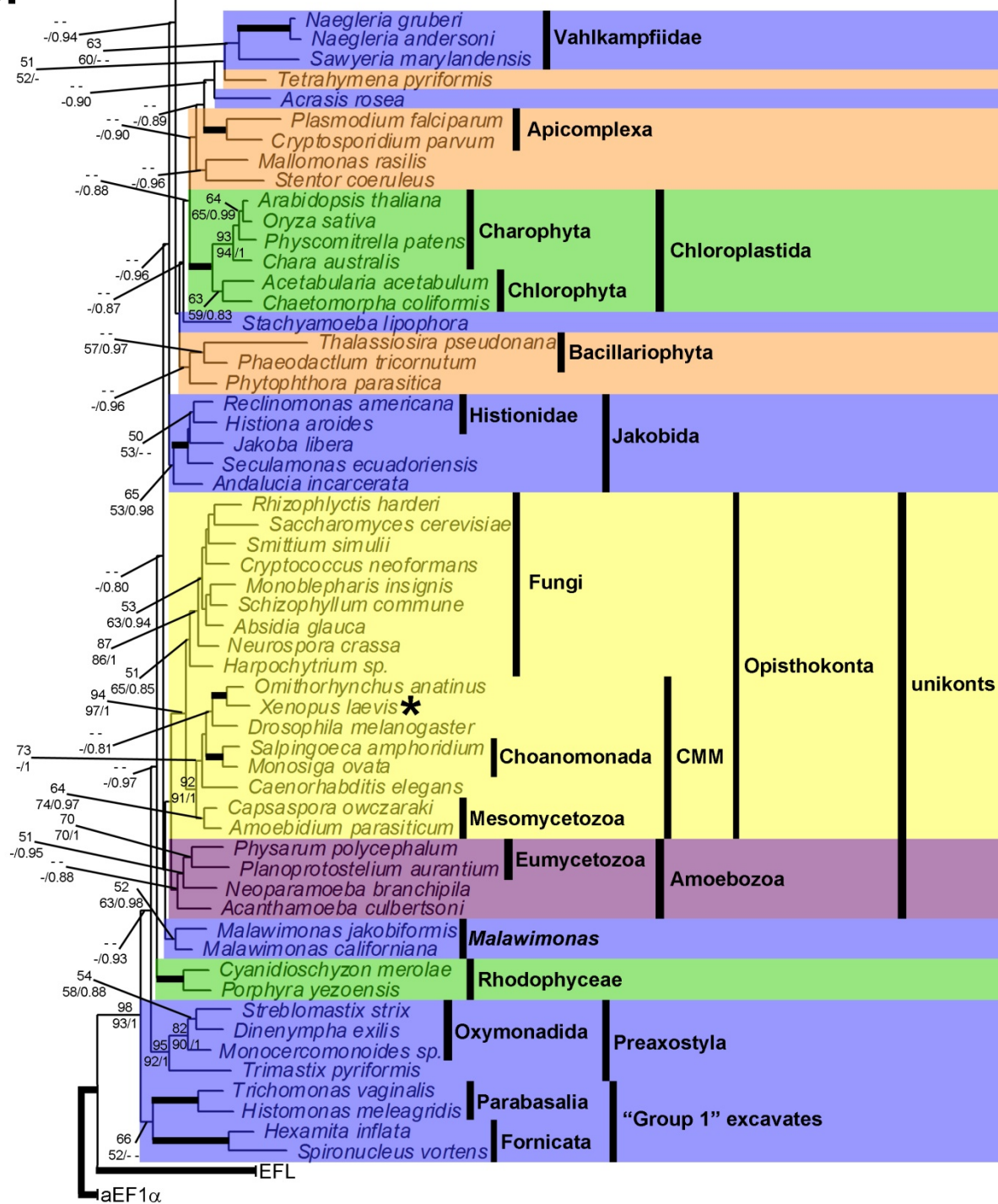


Figure 1.6. The detailed EFL portion of the nucleotide gene tree with its representative OTUs. This is the RAxML ML tree with bootstrap support using the MolcGo3rd alignment with 194 OTUs with the EF1 α and aEF1 α portions of the gene tree summarized. The values on the top are the RAxML bootstrap score before the hash mark and the TNT bootstrap score after; the values below are the GARLI bootstrap score before the hash mark and the MrBayes posterior probability after. Only bootstraps $\geq 50\%$ and posterior probabilities ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bootstrap support from RAxML, TNT, and GARLI as well as ≥ 1.00 posterior probability from the MrBayes resulting tree topologies for the same alignment. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the support thresholds stated. A question mark denotes where the TNT bootstrap score was not recoverable from the Phylogeny.fr website. Colors are indicative of supergroup of the organism in the tree: no color=Archaea, green=Archaeplastida, orange=Chromalveolata, blue=Excavata, yellow=Opisthokonta, and pink=Rhizaria. Groups that are recovered as monophyletic are denoted with vertical bars with the clade name to the right with relatively deeper level clades to the right and relatively tip level clades closer to the OTUs. Cytoplasmic Chlorarachniophyta copy of EFL is labeled with a (c), whereas the EFL copy targeted within the periplastid compartment is labeled with (ppc). The star marks the OTU that still retains the endogenous EF1 α and EFL.

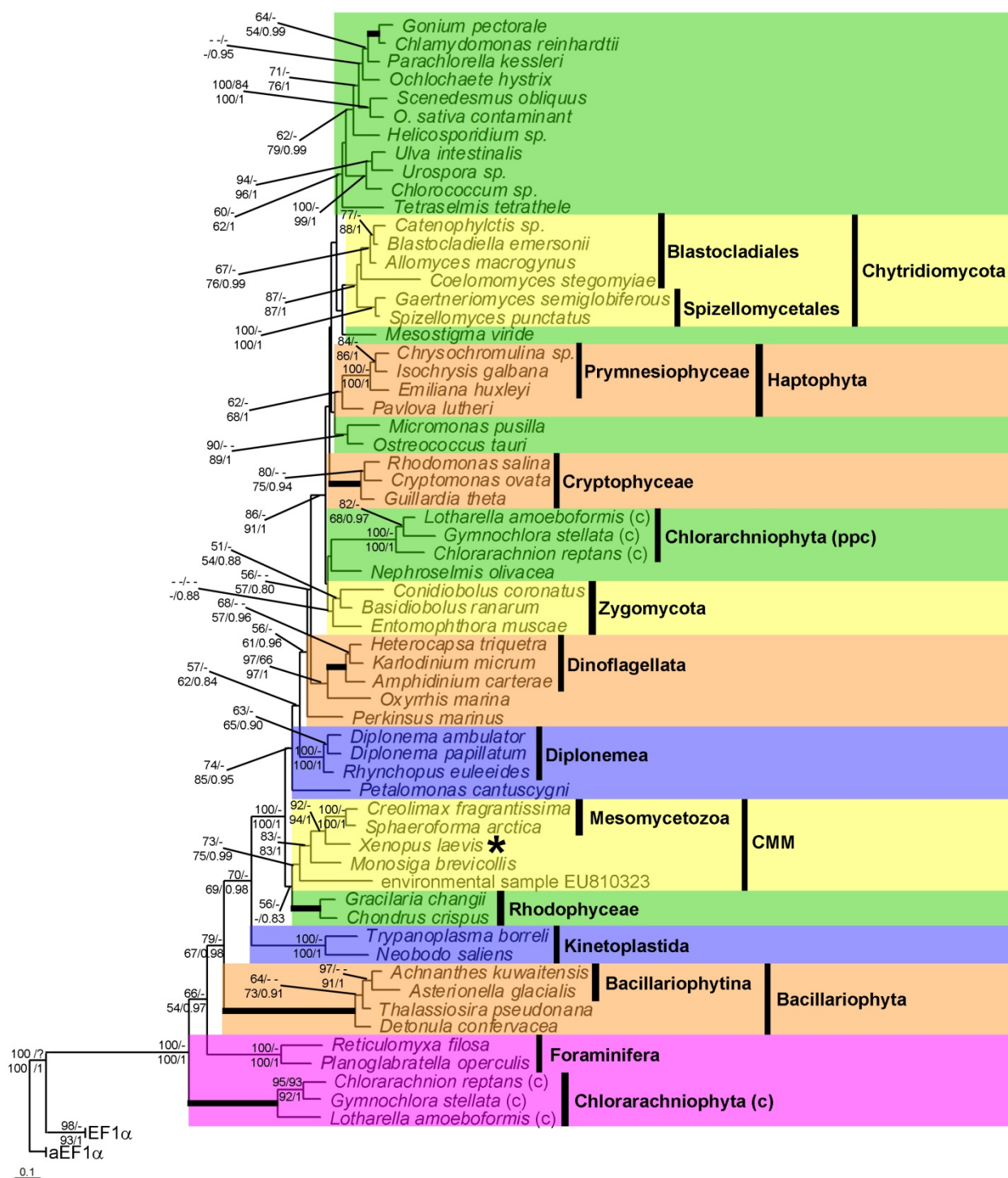


Figure 1.7. The cladogram summary of the RAxML ML amino acid gene tree with clade recovery shown using the MolcGo3rd alignment with 194 OTUs. Recovered relationships are denoted with + at the nodes if both analysis types (RAxML and MrBayes) recover them. Any relationships recovered by only RAxML are not labeled. Gene recovery is denoted the same way, but at the end of the gene name instead of at a node. Colors are indicative of supergroup of the organism in the tree: no color=Archaea, purple=Amoebozoa, green=Archaeplastida, orange=Chromalveolata, blue=Excavata, yellow=Opisthokonta, and pink=Rhizaria.

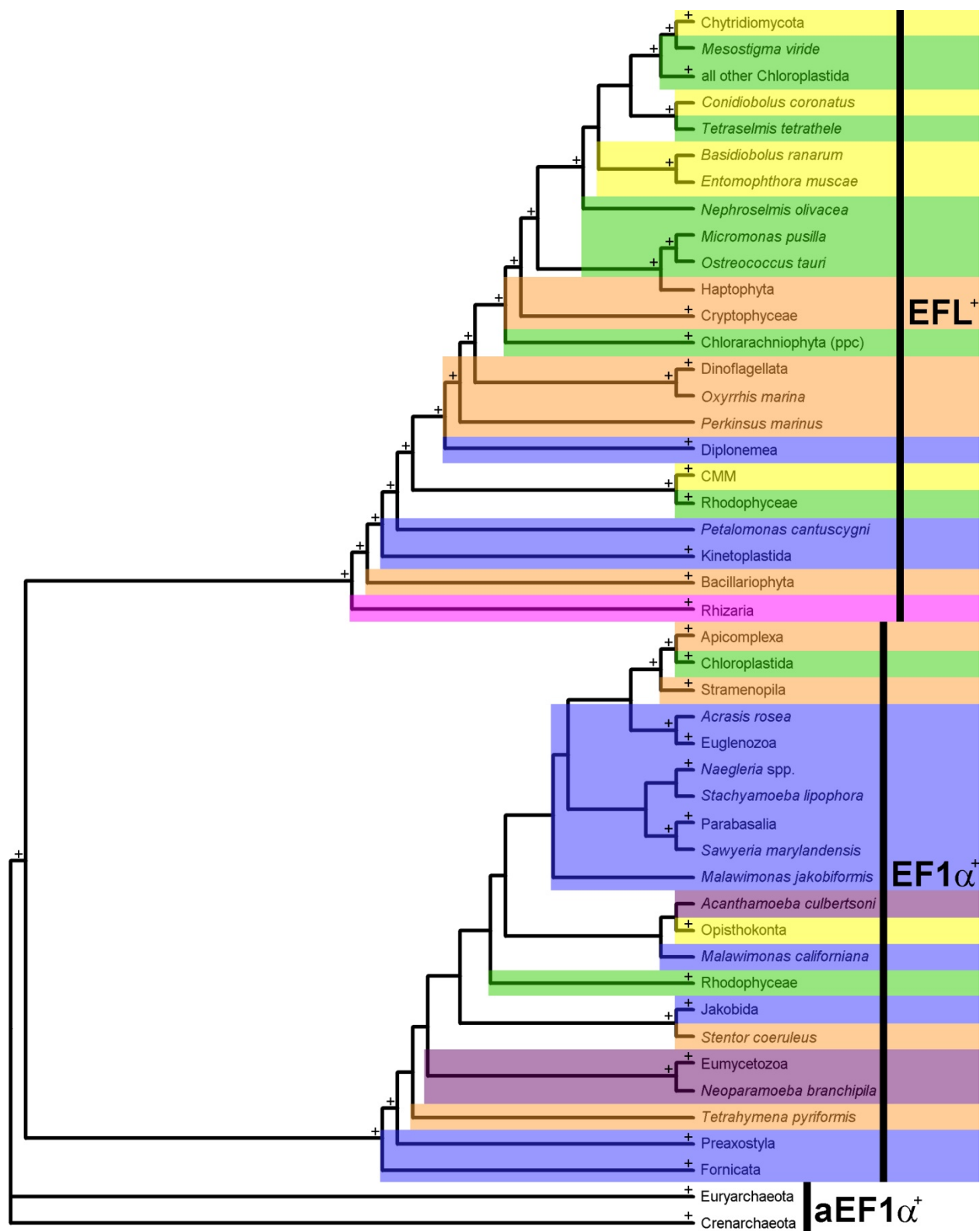
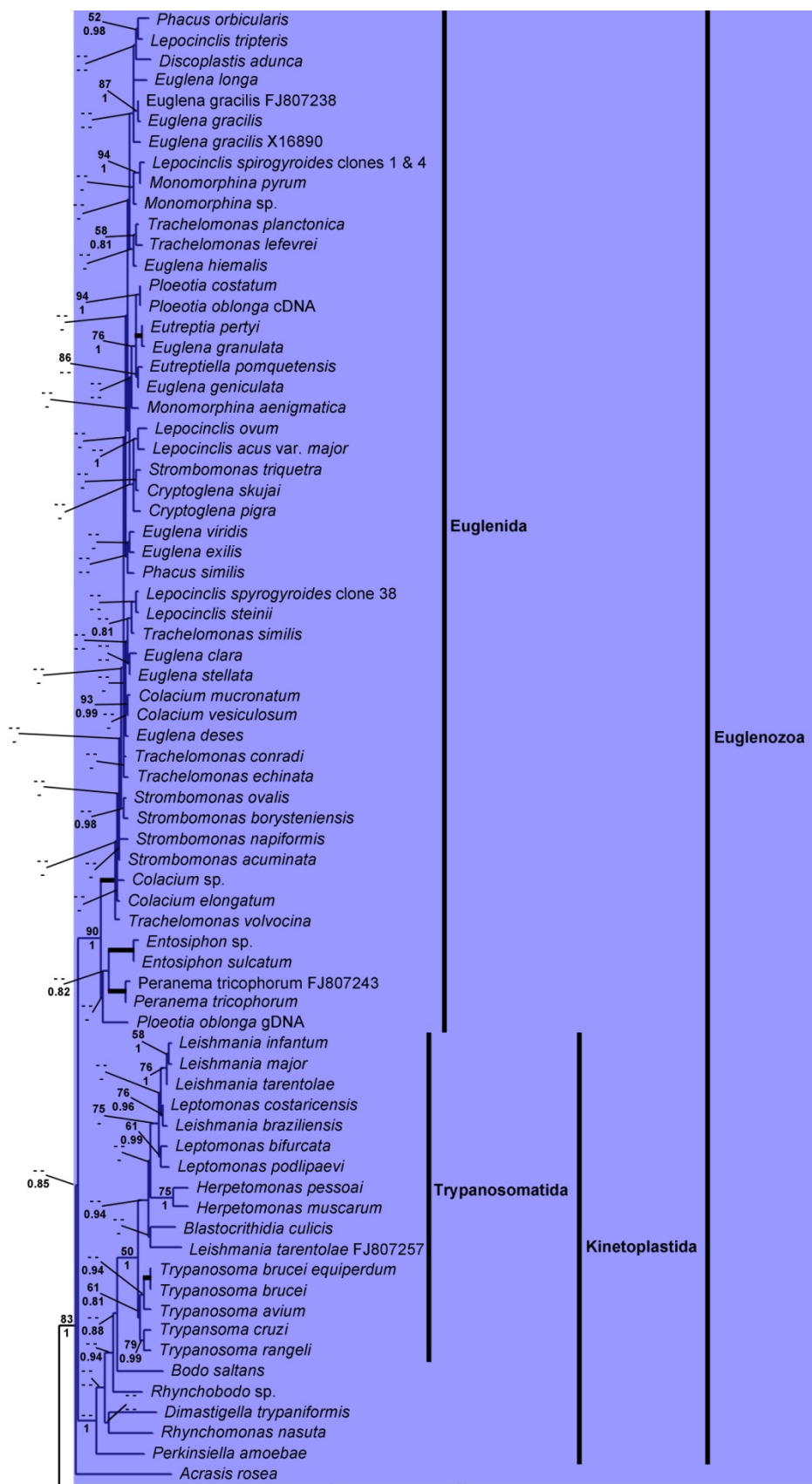


Figure 1.8. The detailed EF1 α portion of the amino acid gene tree with its representative OTUs. This is the RAxML ML tree with bootstrap support using the Mo alignment with 194 OTUs with the EFL and aEF1 α portions of the gene tree summarized. The top value is the RAxML bootstrap score and the MrBayes posterior probability is below. Only bootstraps $\geq 50\%$ and posterior probabilities ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bootstrap support from RAxML and 1.00 posterior probability from the MrBayes resulting tree topologies for the same alignment. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the support thresholds stated. Colors are indicative of supergroup of the organism in the tree: no color=Archaea, purple=Amoebozoa, green=Archaeplastida, orange=Chromalveolata, blue=Excavata and yellow=Opisthokonta. Groups that are recovered as monophyletic are denoted with vertical bars with the clade name to the right with relatively deep level clades to the right and relatively tip level clades closer to the OTUs. The star marks the OTU that still retains the endogenous EF1 α and EFL. **A.** The Euglenozoa and its direct sister OTU portion of the EF1 α tree with a closed black circle indicating where it connects to the rest of the EF1 α detailed tree. **B.** The rest of the eukaryotic EF1 α detailed tree with the closed black circle indicating where it connects to the Euglenozoa OTUs.

A.



B.

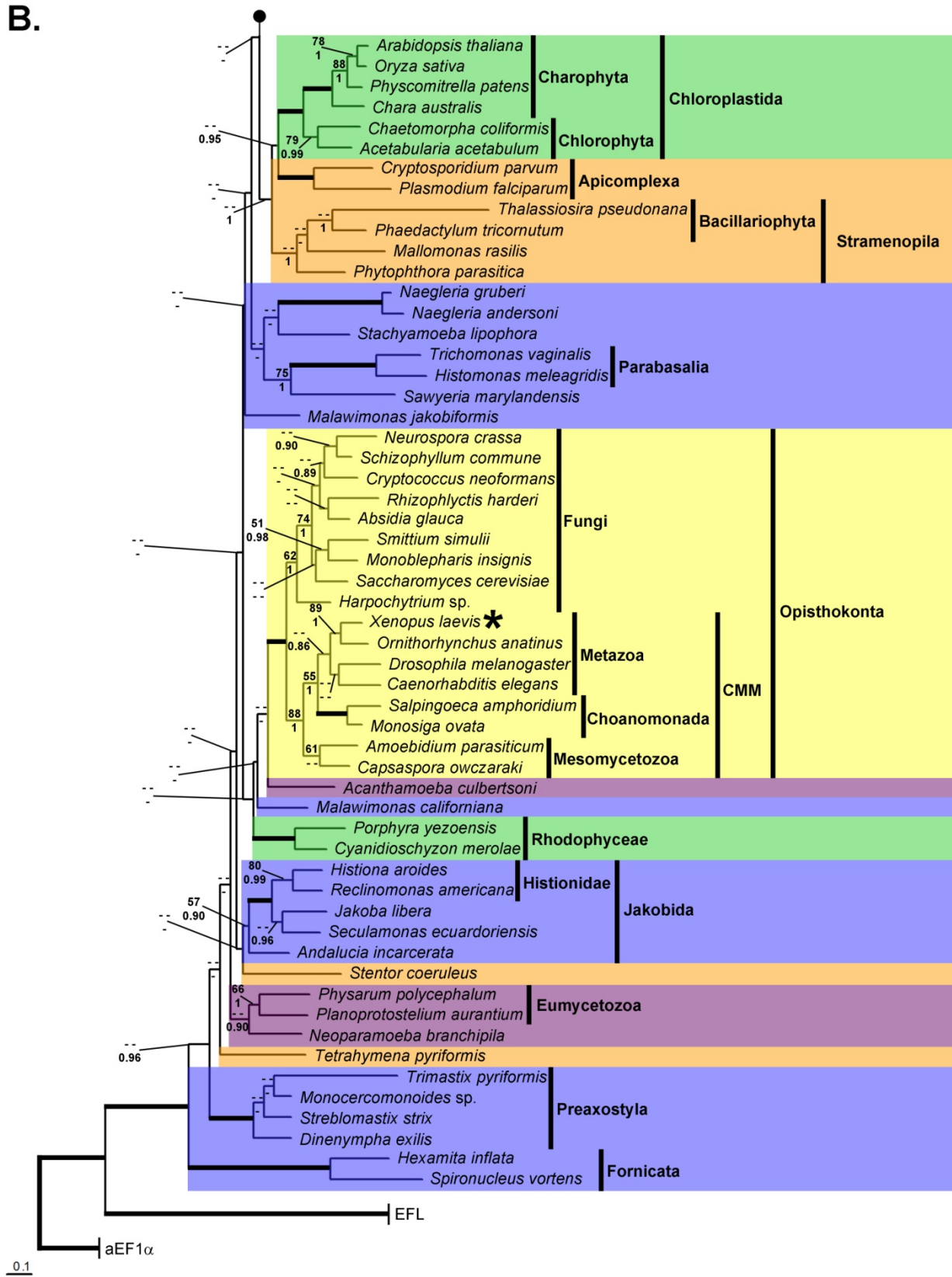


Figure 1.9. The detailed EFL portion of the amino acid gene tree with its representative OTUs. This is the RAxML ML tree with bootstrap support using the Mo alignment with 194 OTUs with the EF1 α and aEF1 α portions of the gene tree summarized. The top values are the RAxML bootstrap scores the MrBayes posterior probabilities are below. Only bootstraps $\geq 50\%$ and posterior probabilities ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bootstrap support from RAxML and ≥ 1.00 posterior probability from the MrBayes resulting tree topologies for the same alignment. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the support thresholds stated. Colors are indicative of supergroup of the organism in the tree: no color=Archaea, green=Archaeplastida, orange=Chromalveolata, blue=Excavata, yellow=Opisthokonta, and pink=Rhizaria. Groups that are recovered as monophyletic are denoted with vertical bars with the clade name to the right with relatively deeper level clades to the right and relatively tip level clades closer to the OTUs. Cytoplasmic Chlorarachniophyta copy of EFL is labeled with a (c), whereas the EFL copy that is targeted within the periplastid compartment is labeled with (ppc). The star marks the OTU that still retains the endogenous EF1 α and EFL.



CHAPTER 2

**EUGLENOZOAN PHYLOGENIES AND EXCAVATE SISTER GROUP ELUCIDATION
WITH EF1 α AND 18S****Abstract:**

Euglenozoans are members of the early-branching Eukaryote lineage known as the Excavata. Euglenozoa is highly supported as a monophyletic clade by both morphological and molecular data, however, the relationships between the three main superclasses within the group, kinetoplastids, diplomonids and euglenids, as well as the sister group to the Euglenozoa remain unresolved. In the past small subunit ribosomal DNA (18S) phylogenies made up the majority of the phylogenies of the Euglenozoa, but protein molecular markers seem to give increased resolution and clarification of the relationships. Protein markers have not been widely used and when they have been (e.g., heat shock proteins, cytochrome oxidase) the analyses have been skewed by the species and outgroup sampling and produced conflicting or ambiguous results. The gene elongation factor 1 alpha (EF1 α) is one such protein coding marker that appears to be very useful in recovering deep-level phylogenetic clades such as those of the Euglenozoa, especially when combined with 18S or other genetic markers. However, it is now apparent that several euglenozoans (the entire diplomonid lineage, some bodonid kinetoplastids and a euglenozoan of unknown affiliation, *Petalomonas cantuscygni* (Euglenozoa *incertae sedis*)), the endogenous EF1 α has been lost either through gene loss or by horizontal gene replacement with the gene elongation factor-like (EFL), which was only described after this research had already begun. Such gene replacement events have occurred from unknown donors on at least two separate occasions. The most parsimonious explanation for the absence of EF1 α is gene loss of

ancient paralogs. Unfortunately, this leaves the sister relationships and basal position of the superclasses unresolved, but the clades within kinetoplastids and euglenids can be resolved as well as the excavate sister lineage to the Euglenozoa since all other known members of the Excavata only express EF1 α and not EFL. We have found that with single gene trees, the relationships between euglenozoan groups are not resolved, but when the genes are analyzed in concatenation, there is resolution and heteroloboseans are strongly supported as the sister lineage to the Euglenozoa. These analyses also recover the Discicristate clade, and the sister to the Discicristata is the Jakobid lineage, which when combined with the Discicristata makes up the Discoba clade.

Introduction:

Euglenozoa

Euglenozoans are an early-branching lineage in the eukaryotic tree of life comprised of unicellular, flagellated protists with three distinctive superclasses and several unassigned genera. The three major Euglenozoa superclasses are the Kinetoplastida, Diplonemea and Euglenida (Cavalier-Smith 1981; Simpson 1997). Kinetoplastids have free-living heterotrophic species, but they also have many parasitic species, including several that cause important human diseases such as: African sleeping sickness (*Trypanosoma brucei*), Chagas' disease (*Trypanosoma cruzi*), and subcutaneous leishmaniasis (*Leishmania major*) (Lukes et al. 2002). The three species together infect more than 20 million people annually and because of their importance, their genomes have been sequenced and compared (Ghedini et al. 2004; Berriman et al. 2005; El-Sayed et al. 2005a; El-Sayed et al. 2005b; Ivens et al. 2005). Diplonemids are under-studied and under-sampled organisms that appear to be mostly comprised of free-living, marine bacteriotrophs (Moreira, Lopez-Garcia, Rodriguez-Valera 2001; Lara et al. 2009), although some

are animal parasites (Kent et al. 1987). The euglenids contain photosynthetic, heterotrophic and osmotrophic species (Leedale 1967; Triemer, Farmer 2007).

The euglenozoan superclasses are unified by several synapomorphic morphological characters and genetic analyses. Originally, Cavalier-Smith (1981) grouped the euglenozoan superclasses based on shared discoidal mitochondrial cristae, non-tubular flagellar hairs, paraxial (or “paraflagellar”) rods associated with the flagellum and a closed pattern of mitosis with an intranuclear spindle. However, these synapomorphies were revised by Simpson (1997) to include only three main characters: (1) the paraxial rod in one or both of the flagella found parallel with the basal body inside the flagellar membrane, (2) the flagellar root pattern of three asymmetrically arranged microtubular roots, and (3) tubular, thick-walled extrusive organelles that may have been lost or reduced in many derived Euglenozoa. The Euglenozoa have been recovered with strong support in molecular analyses using small subunit ribosomal DNA (18S) sequences (Maslov, Yasuhira, Simpson 1999; Preisfeld et al. 2000; Moreira, Lopez-Garcia, Rodriguez-Valera 2001; Preisfeld et al. 2001; Busse, Preisfeld 2002b; Simpson, Lukes, Roger 2002; Busse, Patterson, Preisfeld 2003; Busse, Preisfeld 2003a; Busse, Preisfeld 2003c; Moreira, Lopez-Garcia, Vickerman 2004; Simpson et al. 2004; von der Heyden et al. 2004).

Although the monophyly of Euglenozoa is firmly established, the past molecular phylogenies and morphological studies have failed to provide unequivocal answers to two key questions: (1) Which superclass is sister to the kinetoplastids?; (2) Which superclass is the earliest-branching within the Euglenozoa? There is a great deal of published competing evidence for sister-relationships between superclasses, although most of it is based on a single nuclear-encoded gene, 18S (Table 2.1). The kinetoplastid-diplonemid sister relationship is supported by several lines of evidence, including 18S, heat shock protein (HSP) genes, a

mitochondrial-encoded gene, cytochrome oxidase subunit I (COI), the similarities in genetic code and organization of the mitochondrial genomes into circles (Maslov, Yasuhira, Simpson 1999; Marande, Lukes, Burger 2005), and the similarities in the feeding apparatus of certain bacteriotrophs from both lineages (Montegut-Felkner, Triemer 1994; Montegut-Felkner, Triemer 1996). The euglenid-kinetoplastid sister relationship is also supported by evidence from 18S analyses as well as morphological similarities of the shape of the feeding and flagellar apparatus, patterns of mitosis and cytoskeletal composition (Triemer, Farmer 1991), and some shared plastid-derived genes (Hannaert et al. 2003; Martin, Borst 2003). However, this last evidence appears to be erroneous with the Kinetoplastid plastid-derived genes supported as being incorporated through horizontal gene transfer (HGT) rather than through a common ancestor with a photosynthetic euglenid (Leander 2004; Rogers, Keeling 2004). The diplonemid-euglenid sister relationship is also supported by 18S, COI, and extended long subunit ribosomal DNA (28S) (Maslov, Yasuhira, Simpson 1999; Marin et al. 2003; Ma 2005).

Some of the discrepancies in the euglenozoan phylogeny can be the result of differences in methods and scope of the studies. It should be noted that in several instances the same study produced different sister relationships within the Euglenozoa superclasses depending on the type of analysis or modifications of the analyses performed (i.e., choice of ingroup and outgroup taxa, correction for G+C content in the gene, etc.). It should also be noted that the scope of these studies is not always the same and this leads to over-emphasis for one or more of the superclasses, which in turn biases the relationships examined by this study. The same types of issues arise when attempting to determine the earliest-branching euglenozoan superclass. There are separate instances of each of the three superclasses being supported as basal in different

studies. This conflicting evidence, mostly from the same gene (18S), only confuses the true relationships within the Euglenozoa.

The relationships within the Euglenozoa could be robustly resolved if appropriate markers are used. It is argued that the lack of resolution in the Euglenozoa from the analyses of 18S results from high levels of saturation. Allegedly, there is little phylogenetic signal left in this molecular marker (Busse, Preisfeld 2002a; Busse, Preisfeld 2003a). Also, the sequence for 18S is relatively short with fewer phylogenetically informative sites and is prone to long branch attraction (LBA) artifacts, which lead to strongly supported yet incorrect relationships of taxa (Huelsenbeck 1997). Also, it is very important to have a variety of lines of evidence from “independent” sources to be sure we are studying the overall organismal phylogeny and not the gene or organellar evolutionary history within the organism. Other studies have found that slowly evolving protein sequences (e.g., HSPs) in combination with one or more other data sets can greatly improve phylogenetic resolution (Baldauf, Doolittle 1997; Baldauf et al. 2000; Fast et al. 2002; Stechmann, Cavalier-Smith 2003; Simpson, Roger 2004). However, some of the proteins are not appropriate: COI is too rapidly evolving to resolve the questions asked in this study and HSPs used in other studies are members of complex gene families that could lead to spurious phylogenies if paralogs are used instead of orthologs (Kapoor, Curle, Runham 1995; Sung, Kaplan, Guy 2001; Stechmann, Cavalier-Smith 2003).

EF1 α

EF1 α is an appropriate and useful gene for resolving the euglenozoan phylogeny. EF1 α is a nuclear-encoded slowly-evolving protein that has been successfully used in other studies to resolve deep level phylogenies (Baldauf 2003; Tanabe et al. 2004). Because it is a protein, both amino acid and DNA sequence analyses can be performed with this marker. It can also be

concatenated with previously used markers. EF1 α is robust in phylogenies that correlate with species trees and its use to study the phylogeny of the Euglenozoa is greatly needed due to the potential flaws associated with morphology and the other genetic markers used. However, in the pursuit of acquiring EF1 α sequences from the members of the three superclasses of the Euglenozoa, we have discovered that some bodonid kinetoplastids and all studied members of the diplonemid lineage no longer possess their endogenous EF1 α sequence and the elongation factor-like gene has replaced its functional role in those organisms (Keeling, Inagaki 2004; Gile et al. 2009). Since this research was begun prior to this data obstruction via molecular evolution of genes, we were unable to resolve the relationships within the Euglenozoan superclasses but we were able to utilize EF1 α in concatenation with 18S to clarify the relationships within the kinetoplastid and euglenid lineages as well as to determine the Excavata sister lineages that are closest to the Euglenozoa. Concatenating genes for phylogenetic analyses generally improves the resolution of the lineages because underlying hidden signal from both data sets can be detected when they are together, but not when they are separated (de Queiroz, Gatesy 2007).

The Euglenozoa were historically grouped with the heteroloboseans in a group known as the Discicristates (Cavalier-Smith 1998; Cavalier-Smith 2003); however, recent molecular phylogenetic evidence has shown a closer relationship between Jakobida and Euglenozoa to the exclusion of Heterolobosea (Simpson, Inagaki, Roger 2006). The Discicristates have the synapomorphy of discoidal mitochondrial cristae, but mitochondrial cristae as a phylogenetic character has been shown to be a plastic character in trypanosomatid developmental stages (Brown, Evans, Vickerman 1973) and therefore discoidal mitochondrial cristae may not be an accurate indicator of relatedness between the two lineages. The Discicristates are recovered using molecular phylogenetic analyses as well (Baldauf 2003; Busse, Preisfeld 2003b; von der

Heyden et al. 2004; Ma 2005), but these are the same studies that have relied mostly upon 18S and have limited or uneven in-group sampling, and very questionable outgroup samplings when heteroloboseans are not forced to be the root in the euglenozoan studies. The jakobids are also usually recovered as closely related to the Euglenozoa and together with the heteroloboseans they form a clade recently termed the Discoba (Hampl et al. 2009). The relationships of the jakobids and heteroloboseans as outgroup to the Euglenozoa change depending upon the outgroup taxa sampled as well as whether or not long-branch species are included in the study (Breglia, Slamovits, Leander 2007; Yubuki et al. 2009).

Support and clarification of these relationships between the Euglenozoa and its possible excavate sister groups were unresolved with single gene analyses of EF1 α and 18S because of conflicting topologies or non-monophyly of groups, however, the best resolution of well-supported relationships in this study occurs when EF1 α and 18S are concatenated.

Concatenation results in the recovery of Discoba with the monophyletic Jakobida sister to the Discicristates, comprised of the monophyletic Euglenozoa that still retain EF1 α and the monophyletic Heterolobosea. The relationships of some euglenozoan ingroup tip-level clades are also recovered in the concatenation analyses when neither single gene analysis recovers them. Multigene and various genetic lines of evidence are essential and necessary for improving the deep-level phylogenies for groups such as the Euglenozoa, as is supported by the results from concatenating EF1 α and 18S in this study.

Materials and Methods:

Culture information and nucleic acids extractions

All taxa with sources and accession numbers used in this study are shown in Table 2.2. Taxa were either obtained from culture collections or from colleagues (Triemer et al. 2006). All isolates were grown in the conditions and with the media specified in Chapter 1 of this thesis.

Sequence retrieval, nucleotide isolation, amplification, and sequencing

EF1 α and 18S genes were obtained from euglenids, kinetoplastids, diplomonads, heteroloboseans, jakobids and representatives for all other main Excavata clades with PCR or searches of EST and genomic databases (TBestDB, TriTrypDB v.1.1) or GenBank release 172 (Table 2.2). New sequences were deposited in GenBank with accession numbers XXXX-XXXX: euglenid (1), kinetoplastid (1), jakobid (1), heterolobosean (1).

Total genomic DNA and/or total RNA was extracted from uniprotistan cultures and amplified and sequenced as specified in Chapter 1 of this thesis. The 18S primers and PCR amplification protocols used were based on previously published research (Brosnan et al. 2003; Triemer et al. 2006).

Sequence Alignment

Initially, all 101 EF1 α nucleotide sequences were aligned with MUSCLE (Edgar 2004) and then edited by eye against closely related organisms' EF1 α sequences based on amino acid conceptual translations in Geneious Pro v3.8.5 (Biomatters Ltd.; available through www.geneious.com) (Drummond et al. 2009) and previously published EF1 α alignments (Baldauf, Palmer 1993; Baldauf, Doolittle 1997; Keeling, Inagaki 2004; Fiore-Donno et al. 2005; Gile, Patron, Keeling 2006; Gile, Keeling 2008).

Several different alignments were created from the original alignment by decreasing the number of sequences. In the first culling step, 92 sequences remained and another nine were removed for one or more of the following reasons: (1) The sequence was erroneously placed in

the resulting species tree probably due to either contamination or misidentification of organism, (2) Two or more representatives of the same species clustered together even if there was <95% sequence identity between the sequences, or because (3) The species is not represented by 18S and cannot be used in the concatenated gene studies of EF1 α and 18S (Tables 2.2 and 2.3). The second culling involved removing all eight unikont sequences (Amoebozoa and Opisthokonta), such that only organisms in Excavata remained in the analysis with 84 operational taxonomic units (OTUs).

The 18S gene was acquired for the 92 species represented in the EF1 α alignment and they were combined into a FASTA-formatted document which was aligned with SINA (SILVA Incremental Aligner) on the Silva release 100 comprehensive ribosomal RNA databases website (www.arb-silva.de/) to more accurately align the sequences based on the secondary structure database (Pruesse et al. 2007). This process was repeated for the culled 84 FASTA-formatted file of 18S OTUs.

All resulting gapped alignments from EF1 α and 18S were run through GBlocks v0.91b (Castresana 2000; Talavera, Castresana 2007) on the online server (http://molevol.cmima.csic.es/castresana/Gblocks_server.html) to create reproducible alignments with unsupported residues eliminated (see details in Figs. 1.1, 2.1 and 2.2 and Table 2.3). The GBlocks parameters selected were as follows: codon input data, allowed for smaller blocks, allowed for gap positions within final blocks and allowed for less strict flanking positions. Third codon nucleotides were removed from all alignments with EF1 α to eliminate any homoplastic noise associated with that position to create new alignments: 1a3rd-101, 1a3rd-92, 1a3rd-84, respectively. Concatenation alignments of EF1 α and 18S were created such that the GBlocks alignments for each were put end to end, resulting in 1a18S-92, 1a3rd18S-92, 1a18S-84 and

1a3rd18S-84. Saturation plots were made using the empirical alignment data or the RAxML estimated parameters (see details below for RAxML analyses) to compare each original alignment with all nucleotides included versus its respective alignment with the third codon positions from EF1 α removed based on pairwise likelihood distances calculated in PAUP* (Fig. 2.3) (Swofford 1993).

Phylogenetic analysis

All alignments were processed with Maximum likelihood (ML), and Maximum likelihood estimation using Bayesian algorithms (MLE). ML was performed using RAxML v7.0.4 with simultaneous autodetect rapid bootstrapping (Stamatakis 2006; Stamatakis, Hoover, Rougemont 2008) on the CIPRES server (Miller et al. 2009). MLE was also performed with MrBayes v3.1.2 (Huelsenbeck, Ronquist 2001; Ronquist, Huelsenbeck 2003) on either the CIPRES server or the University of Georgia (UGA) rCluster. For the ML and MLE analyses, the model of sequence evolution (MoSE) was the one specified from the AIC indices from MODELTEST v3.7 (Posada, Crandall 1998; Posada, Crandall 2001; Posada, Buckley 2004). Trees were rooted using either *Acanthamoeba culbertsoni* (Amoebozoan; 101 and 92 OTU alignments) or *Hexamita inflata* (Excavata/Fornicata; 84 OTU alignments). To ensure convergence upon the correct topology MLE consensus using the EF1 α single-gene alignments, four separate runs of thirty million generations each were performed with trees sampled every thousand generations. Of the thirty thousand trees created from each run a conservative fifteen million trees were discarded as burn-in. The resulting fifteen thousand trees from each of the four runs were then combined for a total of sixty thousand trees in PAUP* in a majority rules consensus tree to determine relationships between clades as well as posterior probability (pp) support of the MLE topology. For the 18S and concatenated analyses using MrBayes two sets of

two runs of four million generations each were performed with two million generations discarded as burn-in. Unlike the EF1 α single-gene alignments, these runs were not combined in PAUP*, but were combined within MrBayes itself using the halfcompat setting. Each tree resulting from the combination of two runs (twenty thousand trees in each consensus tree output) were compared for convergence on a single topology. This was achieved at these settings for both the 18S and concatenated gene analyses and the pp values were taken from one of the two consensus trees that were output from MrBayes for each respective analysis. See Tables 2.3 and 2.4 for further details of analyses and MoSE selection for each analysis.

Clade Recovery

All alignments were considered using all analyses types and recovery of expected monophyletic clades based on previous literature were summarized (Table 2.5). A Phylocode-like hierarchical scheme is employed in Table 2.5 to help visualize relative depth in the Eukaryotic branch of the ToL, however, the rank status is not comparable between groups. Clades are represented not only with “supergroups” and “groups” names listed in Table 2.5 (Adl et al. 2005), but more recent clades are also represented that represent larger classification types, such as bikonts, unikonts (Burki, Pawlowski 2006; Burki, Shalchian-Tabrizi, Pawlowski 2008), “Group 1” excavates (Fornicata+Parabasalia), “Group 2” excavates (*Malawimonas*+*Preaxostyla*) and the “Group 3” excavates, now known as the Discoba (Hampl et al. 2009) clades (Simpson, Inagaki, Roger 2006). The currently un-named, but historically robustly supported clade consisting of the Choanomonada, Mesomycetozoa and Metazoa group was termed the CMM clade for the purpose of this study. The total number of recovered clades out of the possible clades we expected to be able to recover, if the gene trees match the species tree, were tabulated for each alignment type to analyze which was best at clade recovery. The alignments with 101

OTUs were expected to recover 46 clades, the 92 OTU alignment 42 clades and the 84 OTU alignment 37 clades.

Results:

Not all euglenozoan superclasses retain endogenous EF1 α

As was described in Chapter 1 of this thesis and in Gile et al. (2009), not all euglenozoan organisms retain their endogenous copy of EF1 α . All diplomonids have lost their EF1 α and are only represented in the EFL portion of the gene family tree (Fig. 2.4). The kinetoplastids that have EFL are recovered as a clade consistently, as are the diplomonids, however, they are never recovered together and *Petalomonas cantuscygni* (Euglenozoa *incertae sedis*) alternates branching closer to one or the other depending on the alignment type and analysis type used to study the gene family (results from Ch. 1). Endogenous EF1 α is present in the majority of the kinetoplastids, all the euglenids and every other major clade in the Excavata and was therefore used to resolve clades within the euglenids and kinetoplastids as well as determining the excavate clade sister to the Euglenozoa.

Single gene analyses

Phylogenies constructed from EF1 α alignments with 101 taxa were found to have several instances of sequence placement that were incongruent with respect to the species they represented (Fig. 2.6). *Leishmania tarentolae* (GenBank Accession FJ807257) is never recovered with the other *Leishmania* species, even the other *L. tarentolae* sequence new to this study, but rather it clusters with the *Trypanosoma* species or *Blastocrithidia culicis*. Similarly, *Euglena gracilis* (X16890) groups with species in other genera within the euglenids rather than with the two other *E. gracilis* sequences that were consistently recovered as sister to one another and within or sister to a clade of other *Euglena* species. Similarly, *Lepocinclis spirogyroides*

was represented by two distinct sequences that had <95% sequence identity and clones 1 and 4 consistently branched with either *Phacus* or *Monomorphina* species, whereas clone 38 branched with other *Lepocinclis* species. Also, although the two *Ploeotia* species represented by cDNA (*P. oblonga* and *P. costatum*) were sisters, they were recovered within the Euglenea rather than with the *P. oblonga* gDNA sequence, which was accurately recovered as basal to the euglenids. There were two species of *Entosiphon* represented in the study and they had >95% sequence identity and consistently were sister to one another. There were also two sequences for the same species of *Peranema trichophorum* that were <95% identical in their sequences, but they too were always recovered as sisters. As mentioned in the Methods section, these problematic or redundant sequences were culled in the following single gene alignments. Otherwise, the 1a-101 and 1a3rd-101 were fairly comparable to the other EF1 α single gene trees.

The 92 and 84 EF1 α alignments consistently recovered excavate clades, but were less able to recover tip-level clades. Specifically, the genus-level euglenid and jakobid clades were almost never recovered using EF1 α alone with any outgroup changes or removal of the 3rd codon position (Figs. 2.5B and 2.7). They were able to recover the Euglenozoa, Euglenida, Euglenea, Kinetoplastida, Trypanosomatida, bikont/Excavata, “Group 1” excavates, Preaxostyla, Discoba and unikonts. Conversely, the 18S alignments are relatively superior at recovering tip-clades, but lack resolution of all the deep-level clades. Many of the euglenid, trypanosomatid and jakobid tip-clades that are not recovered with EF1 α phylogenies are recovered with 18S, but Euglenida and Discoba were never recovered with 18S (Figs. 2.5C and 2.8).

Interestingly, none of the single gene phylogenies were able to recover Jakobida or the Discicristates. This is the case for EF1 α because the group that is sister to the Euglenozoa is comprised of the jakobids and heteroloboseans, but because of the interactions between *Jakoba*

bahamensis (Jakobida) and *Stachyamoeba lipophora* (Heterolobosea) generally being sister to one another, neither clade is monophyletic or resolved as sister to the Euglenozoa to the exclusion of the other (Figs. 2.5A-B, 2.6, and 2.7). The 18S single gene phylogenies don't recover the Jakobida or Discicristates because generally, the members of the genus *Andalucia* are recovered as a clade that is sister to the Euglenozoa, but separate from the other members of the Jakobida (Figs. 2.5C and 2.8).

The third codon position in EF1 α has the potential for containing historical information that allows resolution of relationships, especially tip-level clades, but it can also contribute to noise that covers phylogenetic signal as is evident from the saturation plot analyses. When EF1 α is used in single gene analyses, there are many pairwise distance points that are much higher than would be expected between the 1a and the 1a3rd alignments. This inflated value when third codon positions are included in the alignments suggests that the sequences involved are either involved in LBA or they are OTUs which appear in places in our study that do not match the hypothetical species tree proposed by previous literature. When the points in question were analyzed it was determined that they were all in the latter group, (e.g., *Sawyeria marylandensis*, *Hexamita inflata*, *Spironucleus vortens*, *Trichomonas vaginalis*, *Trimastix pyriformis*, *Herpetomonas muscarum*, *Naegleria gruberi*, etc.). These high pairwise distance values and LBA OTUs were still evident in the concatenated alignments, but the joining of 18S with EF1 α seems to have decreased the overall distance from the trend line to effectively normalize the data points.

Concatenated gene analyses

The concatenated genes generally have more recovery of expected clades as opposed to the single gene trees, but they also recover clades that neither gene used in the concatenation

could recover in the single gene trees. For the case of the 1a18S-92 concatenated gene alignment, the novel tip-level clades recovered are *Monomorphina* and *Phacus* within the Euglenia and the novel deep-level clades recovered are Jakobida and Discicristates (Fig. 2.9). The same novel clades are also recovered with the 1a18S-84 concatenated gene alignment with only the Excavata organisms rooting the tree. The Jakobida are not recovered with the 1a3rd18S alignments with either 92 or 84 OTUs, because the *Andalucia* clade is recovered as separate and sister to the other jakobids. However, the Discicristata are still recovered by those alignments even though they were not recovered with alignments from 1a3rd or 18S with either 92 or 84 OTUs. Tip-level clades are also better resolved by the concatenations involving 3rd codon position removal of EF1 α : 1a3rd18S-92 recovers *Leptomonas* (Kinetoplastida), which is never recovered in any other alignment and 1a3rd18S-84 alignment also recovers *Phacus*.

Generalizations about exclusion of 3rd codon positions and different outgroup organisms

Although excluding the 3rd codon position of EF1 α generally results in reduced overall recovered clades (1a3rd-101=20 vs. 1a-101=21; 1a3rd-92=18 vs. 1a-92=20; 1a3rd18S-92=29 vs. 1a18S-92=32; 1a3rd-84=15 vs. 1a-84=15; 1a3rd18S-84=26 vs. 1a18S-84=27), there are different clades recovered than if all the nucleotides were included for EF1 α in the analyses. Clades that are only recovered by EF1 α or the concatenation of EF1 α and 18S with 3rd codon positions removed tend to be deep-level, whereas clades that are only recovered by EF1 α or the concatenation of EF1 α and 18S with all nucleotides are included, tend to be tip-level. When 3rd codon positions are removed, Euglenozoa, Euglenida, “Group 2” excavates, Discoba and Heterolobosea are recovered, but when they are included clades such as *Cryptoglana*, *Monomorphina*, Metakinetoplastina, *Leishmania*, *Trypanosoma*, and *Malawimonas* are recovered. The long-branching obligate parasite Vahlkampfiidae species are sister to *Peranema*

trichophorum (Euglenida) when 3rd codon positions are included, but they are recovered outside of the Euglenozoa with the rest of the Heterolobosea as a sister group in the concatenated gene analyses or clustered with the Jakobida organisms and then sister to Euglenozoa in the single gene EF1 α analyses.

No matter which outgroup or gene sampling is used to recreate the phylogeny, there are several lineages that are recovered for all alignments. These include: Kinetoplastida, Fornicata, Preaxostyla, Oxymonadida and *Naegleria*. Some are relatively tip-level, but others are best described as deep-level clades.

There are differences in clade recovery for the overlapping possible clades when unikonts (92 OTUs) are used to root the tree as opposed to only members of the Excavata (84 OTUs). In the 18S-84 analyses, the “Group 1” and “Group 2” excavate clades are recovered although they are not recovered in the 18S-92 analyses. The tip-level *Andalucia* clade is recovered in the 1a-84 analyses, whereas the deep-level Discoba clade is recovered with the 1a-92 analyses. In the 1a3rd-84 analyses, “Group 2” excavates and the Heterolobosea are recovered, whereas they are not in the 1a3rd-92 analyses. The tip-level *Phacus* clade is recovered in the 1a3rd18S-84 analyses, whereas the Heteronematina (Euglenida) are recovered in the 1a3rd-92 analyses and the *Leptomonas* (Kinetoplastida) clade is recovered in the 1a3rd18S-92 analyses. The only time that the monophyly of the deep-level Euglenida clade is recovered using the concatenated alignments was when the 84 OTUs were used in the 1a3rd18S-84 MrBayes MLE analysis. It was not recovered in the other concatenated analyses because the long-branching sequence of *Entosiphon sulcatum* and sometimes *Ploeotia oblonga* (gDNA) were found as basal to the kinetoplastids or Euglenozoa rather than to the euglenids.

Discussion:

EFL HGT or EF1 α loss among euglenozoans without EF1 α ?

Unfortunately, the loss of endogenous EF1 α in the diplomonids and some other euglenozoan species greatly limits the utility of using this gene to help resolve controversial relationships within the superclasses of the Euglenozoa. Also, EFL is not useful in recovering the relationships of the Euglenozoa since the two superclasses and *Petalomonas cantuscygni* are not recovered consistently or with support with either superclass. The fact that Euglenozoa is never recovered within the EFL portion of the gene family tree could be explained by at least two HGT events or by the retention of EFL and subsequent loss of the endogenous EF1 α paralog and followed by the rapid diversification of the EFL sequences in those various lineages. As has been shown in Chapter 1, the branch leading to the EFL clade is long and so are the branches in the clade itself because EFL seems to have a higher rate of evolution than EF1 α . Therefore, the most parsimonious explanation seems to be differential loss of the ancient paralogs EFL and EF1 α within the Euglenozoa.

Effects of Third Codon Positions and Outgroup Selection on Clade Recovery

Despite the inability to resolve the euglenozoan clades with EF1 α , the gene still has the power to aid in resolving relationships of clades within the individual superclasses, euglenids and kinetoplastids, as well as the excavate sister group to the Euglenozoa. Fewer clades are generally recovered if the third base pair position is excluded, but that is generally because fewer tip-level clades are recovered and since this study focuses on the deep-level clades, removing the 3rd codon position is generally preferable. The recovery of Euglenozoa and Euglenida is especially important and since they are only recovered by EF1 α when the third codon position is excluded, those alignments provide valuable data. Jakobida is another important deep-level clade that is not recovered when the third codon positions are removed, demonstrating that both

types of data are necessary to draw conclusions about the Euglenozoa and their sister group, especially if LBA sequences are contributing to the topology as the saturation plots indicated. The same can be said of using more or fewer outgroup organisms to root the tree. Since the deep-level clades are only recovered while using one or the other type of outgroup, 92 (Discoba) or 84 (Heterolobosea and Euglenida) but not both in the same gene alignment-type, then it is imperative to analyze the Euglenozoa questions with both types of outgrouping.

Multiple genes in concatenation strengthen phylogenies

Single gene phylogenies are not as robust at resolving phylogenies as concatenated gene phylogenies. The EF1 α single gene trees were fairly robust at resolving outgroup clades, but not euglenozoan ingroup clades. Similarly, the 18S gene recovers many clades and especially tip-level clades, however important clades for this study are not resolved in the single gene tree. The concatenation of the two genes seems to bring out hidden signal that neither single gene alone could demonstrate in the resulting phylogenies (de Queiroz, Gatesy 2007). This is evident due to the fact that there are several important clades, such as the Jakobida and the Discicristata that are only recovered with the analyses that used the concatenated alignments. In some cases 18S lends more signal to concatenated gene trees, but EF1 α definitely helps recover some clades in the concatenation studies that were not recovered by the 18S gene analyses. Overall, the concatenated alignments also show improved posterior probability and bootstrap support for the recovered clades in the MLE and ML analyses, respectively.

Conclusion:

EF1 α cannot be utilized to resolve the sister relationships of the Euglenozoa, nor can it be used to help determine which superclass is basal within the Euglenozoa because the diplonemid

superclass seems to have lost its endogenous EF1 α early in its evolution from either the kinetoplastid or euglenid superclass and instead retained its ancient paralog, EFL.

There were several conclusions that could be drawn from this study based on the systematic methods that were considered in the analyses. Third codon position inclusion or exclusion as well as alternative rooting strategies are both required to draw conclusions about the relationships within the Euglenozoa and within the Discoba. There seems to be both historical and homoplasious information in the third codon position since deep-level relationships are recovered in analyses using alignments with all positions that are not recovered in analyses with alignments with the third codon positions removed and vice versa. Also, using alternate rooting strategies with either distantly related outgroups such as the unikonts as opposed to just using the Excavata lineages recovers important deep-level clades differently. Using concatenated EF1 α and 18S genes effectively gives two different lines of evidence to help resolve the Euglenozoa; however, analyzing the phylogenies of the single gene trees also helps to infer where signal for certain clades is coming from, especially when the topologies of the different gene analyses are incongruent with one another.

Questions about the phylogenetic relationships of the Euglenozoa superclasses and about the excavate sister group to the Euglenozoa are resolved using the single genes EF1 α and 18S as well as the concatenated gene analyses. Not only were the Euglenida, Kinetoplastida and Euglenozoa clades recovered, but also the Heterolobosea were recovered as sister to the Euglenozoa to support the Discicristata clade. The contentious relationship between the Jakobida and either the Euglenozoa or the Heterolobosea was elucidated. Jakobida was recovered as sister to the Discicristata supporting the grouping of the Jakobida, Heterolobosea

and Euglenozoa in the Discoba clade. These findings will enable future studies of the Euglenozoa to utilize heteroloboseans and perhaps also jakobids as appropriate outgroup taxa.

References:

- Adl, SM, AGB Simpson, MA Farmer, et al. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology* 52:399-451.
- Baldauf, SL. 2003. The deep roots of eukaryotes. *Science* 300:1703-1706.
- Baldauf, SL, WF Doolittle. 1997. Origin and evolution of the slime molds (Mycetozoa). *Proceedings of the National Academy of Sciences of the United States of America* 94:12007-12012.
- Baldauf, SL, JD Palmer. 1993. Animals and Fungi Are Each Others Closest Relatives - Congruent Evidence from Multiple Proteins. *Proceedings of the National Academy of Sciences of the United States of America* 90:11558-11562.
- Baldauf, SL, AJ Roger, I Wenk-Siefert, WF Doolittle. 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* 290:972-977.
- Berriman, ME Ghedin C Hertz-Fowler, et al. 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309:416-422.
- Breglia, SA, CH Slamovits, BS Leander. 2007. Phylogeny of phagotrophic euglenids (Euglenozoa) as inferred from HSP90 gene sequences. *Journal of Eukaryotic Microbiology* 54:86-92.
- Brosnan, S, W Shin, KM Kjer, RE Triemer. 2003. Phylogeny of the photosynthetic euglenophytes inferred from the nuclear SSU and partial LSU rDNA. *International Journal of Systematic and Evolutionary Microbiology* 53:1175-1186.
- Brown, RC, DA Evans, K Vickerman. 1973. Changes in oxidative metabolism and ultrastructure accompanying differentiation of the mitochondrion in *Trypanosoma brucei*. *International Journal for Parasitology* 3:691-698, IN617-IN622, 699-704.
- Burki, F, J Pawlowski. 2006. Monophyly of Rhizaria and multigene phylogeny of unicellular bikonts. *Molecular Biology and Evolution* 23:1922-1930.
- Burki, F, K Shalchian-Tabrizi, J Pawlowski. 2008. Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. *Biology Letters* 4:366-369.
- Busse, I, DJ Patterson, A Preisfeld. 2003. Phylogeny of phagotrophic euglenids (Euglenozoa): A molecular approach based on culture material and environmental samples. *Journal of Phycology* 39:828-836.
- Busse, I, A Preisfeld. 2002a. Phylogenetic position of *Rhynchopus* sp. and *Diplonema ambulator* as indicated by analyses of euglenozoan small subunit ribosomal DNA. *Gene* 284:83-91.
- Busse, I, A Preisfeld. 2002b. Unusually expanded SSU ribosomal DNA of primary osmotrophic euglenids: Molecular evolution and phylogenetic inference. *Journal of Molecular Evolution* 55:757-767.

- Busse, I, A Preisfeld. 2003a. Application of spectral analysis to examine phylogenetic signal among euglenid SSU rDNA data sets (Euglenozoa). *Organisms Diversity & Evolution* 3:1-12.
- Busse, I, A Preisfeld. 2003b. Discovery of a group I intron in the SSU rDNA of *Ploeotia costata* (Euglenozoa). *Protist* 154:57-69.
- Busse, I, A Preisfeld. 2003c. Systematics of primary osmotrophic euglenids: a molecular approach to the phylogeny of Distigma and Astasia (Euglenozoa). *International Journal of Systematic and Evolutionary Microbiology* 53:617-624.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17:540-552.
- Cavalier-Smith, T. 1981. Eukaryote kingdoms: seven or nine? *Biosystems* 14:461-481.
- Cavalier-Smith, T. 1998. A revised six-kingdom system of life. *Biological Reviews* 73:203-266.
- Cavalier-Smith, T. 2003. The excavate protozoan phyla Metamonada Grasse emend. (Anaeromonadea, Parabasalia, Carpediemonas, Eopharyngia) and Loukozoa emend. (Jakobea, Malawimonas): their evolutionary affinities and new higher taxa. *International Journal of Systematic and Evolutionary Microbiology* 53:1741-1758.
- de Queiroz, A, J Gatesy. 2007. The supermatrix approach to systematics. *Trends in Ecology & Evolution* 22:34-41.
- Drummond, AJ, B Ashton, M Cheung, J Heled, M Kearse, R Moir, S Stones-Havas, T Thierer, A Wilson. 2009. Geneious: Biomatters Ltd.
- Edgar, RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792-1797.
- El-Sayed, N, P Myler, D Bartholomeu, et al. 2005a. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309:409 - 415.
- El-Sayed, N, P Myler, G Blandin, et al. 2005b. Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309:404 - 409.
- Fast, NM, LR Xue, S Bingham, PJ Keeling. 2002. Re-examining alveolate evolution using multiple protein molecular phylogenies. *Journal of Eukaryotic Microbiology* 49:30-37.
- Fiore-Donno, AM, CD Berney, J Pawlowski, SL Baldauf. 2005. Higher-order phylogeny of plasmodial slime molds (Myxogastria) based on elongation factor 1-A and small subunit rRNA gene sequences. *Journal of Eukaryotic Microbiology* 52:201-210.
- Ghedini, E, F Bringaud, J Peterson, et al. 2004. Gene synteny and evolution of genome architecture in trypanosomatids. *Molecular and Biochemical Parasitology* 134:183-191.

- Gile, GH, D Faktorová, CA Castlejohn, G Burger, BF Lang, MA Farmer, J Lukeš, PJ Keeling. 2009. Distribution and Phylogeny of EFL and EF-1 α in Euglenozoa Suggest Ancestral Co-Occurrence Followed by Differential Loss. *PLoS ONE* 4.
- Gile, GH, PJ Keeling. 2008. Nucleus-encoded periplastid-targeted EFL in chlorarachniophytes. *Molecular Biology and Evolution* 25:1967-1977.
- Gile, GH, NJ Patron, PJ Keeling. 2006. EFL GTPase in cryptomonads and the distribution of EFL and EF-1 alpha in chromalveolates. *Protist* 157:435-444.
- Hampl, V, L Hug, JW Leigh, JB Dacks, BF Lang, AGB Simpson, AJ Roger. 2009. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". *Proceedings of the National Academy of Sciences of the United States of America* 106:3859-3864.
- Hannaert, V, E Saavedra, F Duffieux, J Szikora, D Rigden, P Michels, F Opperdoes. 2003. Plant-like traits associated with metabolism of *Trypanosoma* parasites. *Proc Natl Acad Sci USA* 100:1067 - 1071.
- Huelsenbeck, JP. 1997. Is the Felsenstein zone a fly trap? *Systematic Biology* 46:69-74.
- Huelsenbeck, JP, F Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.
- Ivens, ACCS PeacockEA Worthey, et al. 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309:436-442.
- Kapoor, M, CA Curle, C Runham. 1995. The HSP70 gene family of *Neurospora crassa* - cloning, sequence-analysis, expression, and genetic-mapping of the major stress-inducible member. *Journal of Bacteriology* 177:212-221.
- Keeling, PJ, Y Inagaki. 2004. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1 alpha. *Proceedings of the National Academy of Sciences of the United States of America* 101:15380-15385.
- Kent, ML, RA Elston, TA Nerad, TK Sawyer. 1987. An *Isonema*-like flagellate (Protozoa, Mastigophora) infection in larval geoduck clams, *Panope abrupta*. *Journal of Invertebrate Pathology* 50:221-229.
- Lara, E, D Moreira, A Vereshchaka, P Lopez-Garcia. 2009. Pan-oceanic distribution of new highly diverse clades of deep-sea diplomonads. *Environmental Microbiology* 11:47-55.
- Leander, BS. 2004. Did trypanosomatid parasites have photosynthetic ancestors? *Trends in Microbiology* 12:251-258.
- Leedale, G. 1967. Euglenoid Flagellates. 1-73.
- Lukes, J, DL Guilbride, J Votypka, A Zikova, R Benne, PT Englund. 2002. Kinetoplast DNA network: Evolution of an improbable structure. *Eukaryotic Cell* 1:495-502.

- Ma, B. 2005. Phylogeny of deep-level relationships within Euglenozoa based on combined small subunit and large subunit ribosomal DNA sequences. Cellular Biology. Athens: University of Georgia.
- Marande, W, J Lukes, G Burger. 2005. Unique mitochondrial genome structure in diplomemids, the sister group of kinetoplastids. Eukaryotic Cell 4:1137-1146.
- Marin, B, A Palm, M Klingberg, M Melkonian. 2003. Phylogeny and taxonomic revision of plastid-containing euglenophytes based on SSU rDNA sequence comparisons and synapomorphic signatures in the SSU rRNA secondary structure. Protist 154:99-145.
- Martin, W, P Borst. 2003. Secondary loss of chloroplasts in trypanosomes. Proceedings of the National Academy of Sciences of the United States of America 100:765-767.
- Maslov, DA, S Yasuhira, L Simpson. 1999. Phylogenetic affinities of *Diplonema* within the euglenozoa as inferred from the SSU rRNA gene and partial COI protein sequences. Protist 150:33-42.
- Miller, MA, MT Holder, R Vos, PE Midford, T Liebowitz, L Chan, P Hoover, T Warnow. 2009. The CIPRES Portals.
- Montegut-Felkner, AE, RE Triemer. 1994. Phylogeny of *Diplonema ambulator* (Larsen and Patterson). 1. Homologies of the flagellar apparatus. European Journal of Protistology 30:227-237.
- Montegut-Felkner, AE, RE Triemer. 1996. Phylogeny of *Diplonema ambulator* (Larsen and Patterson) .2. Homologies of the feeding apparatus. European Journal of Protistology 32:64-76.
- Moreira, D, P Lopez-Garcia, F Rodriguez-Valera. 2001. New insights into the phylogenetic position of diplomemids: G+C content bias, differences of evolutionary rate and a new environmental sequence. International Journal of Systematic and Evolutionary Microbiology 51:2211-2219.
- Moreira, D, P Lopez-Garcia, K Vickerman. 2004. An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: proposal for a new classification of the class Kinetoplastea. International Journal of Systematic and Evolutionary Microbiology 54:1861-1875.
- Posada, D, TR Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic Biology 53:793-808.
- Posada, D, KA Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.
- Posada, D, KA Crandall. 2001. Selecting the best-fit model of nucleotide substitution. Systematic Biology 50:580-601.

- Preisfeld, A, S Berger, I Busse, S Liller, HG Ruppel. 2000. Phylogenetic analyses of various euglenoid taxa (Euglenozoa) based on 18S rDNA sequence data. *Journal of Phycology* 36:220-226.
- Preisfeld, A, I Busse, M Klingberg, S Talke, HG Ruppel. 2001. Phylogenetic position and inter-relationships of the osmotrophic euglenids based on SSU rDNA data, with emphasis on the Rhabdomonadales (Euglenozoa). *International Journal of Systematic and Evolutionary Microbiology* 51:751-758.
- Pruesse, E, C Quast, K Knittel, BM Fuchs, W Ludwig, J Peplies, FO Glockner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucl. Acids Res.* 35:7188-7196.
- Rogers, M, PJ Keeling. 2004. Lateral transfer and recompartmentalization of Calvin cycle enzymes of plants and algae. *Journal of Molecular Evolution* 58:367-375.
- Ronquist, F, JP Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
- Simpson, A, J Lukes, A Roger. 2002. The evolutionary history of kinetoplastids and their kinetoplasts. *Molecular Biology and Evolution* 19:2071 - 2083.
- Simpson, AGB. 1997. The identity and composition of the Euglenozoa. 148:318-328.
- Simpson, AGB, EE Gill, HA Callahan, RW Litaker, AJ Roger. 2004. Early evolution within kinetoplastids (Euglenozoa), and the late emergence of trypanosomatids. *Protist* 155:407-422.
- Simpson, AGB, Y Inagaki, AJ Roger. 2006. Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of "primitive" eukaryotes. *Molecular Biology and Evolution* 23:615-625.
- Simpson, AGB, AJ Roger. 2004. Protein phylogenies robustly resolve the deep-level relationships within Euglenozoa. *Molecular Phylogenetics and Evolution* 30:201-212.
- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.
- Stamatakis, A, P Hoover, J Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Systematic Biology* 57:758-771.
- Stechmann, A, T Cavalier-Smith. 2003. Phylogenetic analysis of eukaryotes using heat-shock protein HSP90. *Journal of Molecular Evolution* 57:408-419.
- Sung, DY, F Kaplan, CL Guy. 2001. Plant HSP70 molecular chaperones: Protein structure, gene family, expression and function. *Physiologia Plantarum* 113:443-451.
- Swofford, DL. 1993. PAUP - A computer program for phylogenetic inference using Maximum Parsimony. *Journal of General Physiology* 102:A9-A9.

- Talavera, G, J Castresana. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. 56:564-577.
- Tanabe, Y, M Saikawa, MM Watanabe, J Sugiyama. 2004. Molecular phylogeny of Zygomycota based on EF-1 alpha and RPB1 sequences: limitations and utility of alternative markers to rDNA. Molecular Phylogenetics and Evolution 30:438-449.
- Triemer, RE, MA Farmer. 1991. An ultrastructural comparison of the mitotic apparatus, feeding apparatus, flagellar apparatus and cytoskeleton in euglenoids and kinetoplastids. Protoplasma 164:91-104.
- Triemer, RE, MA Farmer. 2007. A decade of euglenoid molecular phylogenetics. In: J Brodie, J Lewis, editors. Unraveling the algae: the past, present and future of algal systematics: Systematics Association Series, CRC Press. p. 315-330.
- Triemer, RE, E Linton, W Shin, A Nudelman, A Monfils, M Bennett, S Brosnan. 2006. Phylogeny of the euglenales based upon combined SSU and LSU rDNA sequence comparisons and description of *Discoplastis* gen. nov (Euglenophyta). Journal of Phycology 42:731-740.
- von der Heyden, S, EE Chao, K Vickerman, T Cavalier-Smith. 2004. Ribosomal RNA phylogeny of bodonid and diplomid flagellates and the evolution of euglenozoa. Journal of Eukaryotic Microbiology 51:402-416.
- Yubuki, N, VP Edgcomb, JM Bernhard, BS Leander. 2009. Ultrastructure and molecular phylogeny of *Calkinsia aureus*: cellular identity of a novel clade of deep-sea euglenozoans with epibiotic bacteria. BMC Microbiology 9.

Table 2.1. Summary of the phylogenetic controversy of the euglenozoan lineages. Table based on literature that focused on at least one euglenozoan superclass or euglenozoan *incertae sedis* organism and had at least one representative from the three superclasses in the analysis. +=euglenozoan Relationship supported in study (two listed superclasses related as sisters to the exclusion of the third superclass); Basal/Out.= basal Euglenozoan superclass/outgroup organisms used in study; K=Kinetoplastida; D=Diplonemea; E=Euglenida; H=Heterolobosea; J=Jakobida; euk=other eukaryotes in several main branches of the eukaryotic ToL; p=polyphyletic superclass; COI = cytochrome c oxidase subunit I (mitochondrial); 18S = small subunit ribosomal DNA; HSP = heat shock protein (number is size in kilodaltons; indicative of HSP gene family); 28S = large subunit ribosomal DNA; ML=maximum likelihood; MP=maximum parsimony; NJ=neighbor joining distance method; ‡=References: 1=(Maslov, Yasuhira, Simpson 1999); 2=(Moreira, Lopez-Garcia, Rodriguez-Valera 2001); 3=(Preisfeld et al. 2001); 4=(Busse, Preisfeld 2002b); 5=(Simpson, Lukes, Roger 2002); 6=(Busse, Preisfeld 2003c); 7=(Busse, Patterson, Preisfeld 2003); 8=(Moreira, Lopez-Garcia, Vickerman 2004); 9=(Simpson et al. 2004); 10=(Simpson, Roger 2004); 11=(Breglia, Slamovits, Leander 2007); 12=(Busse, Preisfeld 2002a); 13=(Busse, Preisfeld 2003b); 14=(Busse, Preisfeld 2003a); 15=(Marin et al. 2003); 16=(von der Heyden et al. 2004); 17=(Ma 2005); 18=(Yubuki et al. 2009); †=the same study produced different sister relationships within the euglenozoan superclasses depending upon the type of analysis or modifications of the analyses performed.

Rel+	Basal/Out.	Data	Method(s)	Emphasis	Ref ‡
KD†	E; euk	COI (& its genetic code) & 18S	ML tree (also ML consensus tree, MP & NJ, but supported DE)	D in Euglenozoa	1
KD†	E; euk ML (high G+C)	18S	LogDet, NJ, MP, ML (also varied high or low G+C outgroups)	KDE in Euglenozoa	2
KD	E; euk	18S	ML, MP, NJ	osmotrophic E	3
KD	E; euk	18S (& its 2° structure)	ML	osmotrophic E	4
KD	E; E only	18S & HSP90	ML & MLdistance	K	5
KD	E; euk	18S	ML & MP	osmotrophic E	6
KD	pE; euk	18S	ML, MP, Bayesian & spectral analysis	phagotrophic E	7
KD†	E; euk	18S	NJ, MP, ML & Bayesian	D in Euglenozoa	8
KD	E; euk(70 & combined) E; H+euk(90)	HSP70, HSP90, combined	ML & MLdistance	K in Euglenozoa	9
KD	E; euk+H(70) E; H+euk(90 & combined)	HSP70, HSP90, combined	ML & MLdistance	KDE in Euglenozoa	10
KD	E; J pE; J,H,plants	HSP90	ML, NJ & Bayesian (in second analysis E was polyphyletic)	phagotrophic E	11
EK	D; euk	18S	ML & MP	D in Euglenozoa	12
EK	pE; H+euk	18S	ML, MP, NJ and some 18S 2° structure	phagotrophic E	13

			[(someE,(D,(mostE,K)))]		
EK	D; J+euk+H pE; J	18S	ML [in second analysis E was polyphyletic; (someE,(D,(mostE,K)))]	<i>Calkinsia</i> in Euglenozoa; KD or E	18
DE†	E; euk	COI (& its genetic code) & 18S	ML consensus tree, MP & NJ (also ML tree, but supported KD)	D in Euglenozoa	1
DE†	K; euk ML & LogDet (low G+C) & LogDet (high G+C)	18S	LogDet, NJ, MP, ML(also varied high or low G+C outgroups)	KDE in Euglenozoa	2
DE	K; euk	18S	NJ, MP, ML & spectral analysis	osmotrophic E	14
DE	K; K only	18S	NJ, MP & ML	photosynthetic E	15
DE†	K; K only	18S	NJ, MP, ML & Bayesian	D in Euglenozoa	8
DE	K, H	18S	BioNJ, NJ & MP	K & D in Euglenozoa	16
DE	K, H+euk	18S & extended 28S	NJ, MP, ML, Bayesian	<i>Petalomonas</i> in Euglenozoa; KD or E	17

Table 2.2. Organism information for sequences used in this study. A check mark in gene column denotes whether the sequence is EFL or EF1 α ; a check mark in the alignment columns represents the inclusion of that OTU in either the 101, 92 or 84 OTU alignments, with “R” specifying which OTU was forced as the root of the analyses using that alignment type; bold denotes new sequences obtained in this study; a dash in strain/isolate/clone column represents that that information was not available from the database; “Triemer” represents the sequence was acquired from the Triemer lab at Michigan State University (MSU); and an asterisk denotes different strain used for one or more of the genes in the concatenation study. A dash indicates that 18S was not used or not available in this study. Supergroup and “group” names were from (Adl et al. 2005) and updates from recent literature.

Supergroup	“Group(s)”	Gene		Alignment			Short Name	Species	Strain/ Isolate/Clone	EF1 α Database Accession #	18S Database Accession #
		EF1 α	EFL	101	92	84					
Amoebozoa	Acanthamoebidae	√		√; R	√; R		ACACU	<i>Acanthamoeba culbertsoni</i>	-	AY582829; gDNA	AF019067
Excavata	Euglenozoa/Diplonemea/ <i>Diplonema</i>		√				-	<i>Diplonema ambulator</i>	ATCC 50223	FJ807247; mRNA	-
	Euglenozoa/Diplonemea/ <i>Diplonema</i>		√				-	<i>Diplonema papillatum</i>	ATCC 50162	FJ807248; mRNA	-
	Euglenozoa/Diplonemea		√				-	<i>Rhynchopus euleeides</i>	ATCC 50226	FJ807249; mRNA	-
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	CEL	<i>Colacium elongatum</i>	M11	apply for; cDNA	Triemer
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	MUC	<i>Colacium mucronatum</i>	UTEX LB 2524	apply for; cDNA	AF326232
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	CUA	<i>Colacium</i> sp.	NJ	apply for; cDNA	Triemer
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	CVL	<i>Colacium vesiculosum</i>	UW Lazienki	apply for; cDNA	DQ249874
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	CPI	<i>Cryptoglena pigra</i>	CCAP 1212/1	apply for; cDNA	AJ532437
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	CSK	<i>Cryptoglena skujai</i>	SAG 10.88	apply for; cDNA	AJ532438
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	DAA	<i>Discoplastis adunca</i>	ASW 08039	apply for; cDNA	AJ532453
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	ECL	<i>Euglena clara</i>	SAG 25.98	apply for; cDNA	AJ532423
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	EDS	<i>Euglena deses</i>	SAG 1224-23	apply for; cDNA	AJ532407
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	EXS	<i>Euglena exilis</i>	SAG 1224-11a	apply for; cDNA	DQ140147
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	EGN	<i>Euglena geniculata</i>	SAG 1224-21	apply for; cDNA	AY523034
	Euglenozoa/Euglenida/Euglenea	√		√			EUGGR	<i>Euglena gracilis</i>	Indiana University Z strain	X16890; mRNA	-
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	GRC	<i>Euglena gracilis</i>	Z strain	apply for; cDNA	AF283308
	Euglenozoa/Euglenida/Euglenea	√		√			EUGGG	<i>Euglena gracilis</i>	Z strain	FJ807238; mRNA	-
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	EGR	<i>Euglena granulata</i>	OR5	apply for; cDNA	AJ532422
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	EHS	<i>Euglena hiemalis</i>	CCAP 1224/35	apply for; cDNA	DQ140157
	Euglenozoa/Euglenida/Euglenea	√		√	√	√	EUGLO	<i>Euglena</i> (aka <i>Astasia</i>) <i>longa</i>	- ; SAG 1204-17a*	FJ807237; mRNA	AJ532428*

	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	ESL	<i>Euglena stellata</i>	IAM E-11	apply for; cDNA	AF090868
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	EVI	<i>Euglena viridis</i>	NJ001	apply for; cDNA	AY523036
	Euglenozoa/Euglenida/Euglenea/ Eutreptiales	✓		✓	✓	✓	EPE	<i>Eutreptia pertyi</i>	UTEX LB 1290	apply for; cDNA	AF081589
	Euglenozoa/Euglenida/Euglenea/ Eutreptiales	✓		✓	✓	✓	TET	<i>Eutreptiella pomquetensis</i>	CCMP 1491	apply for; cDNA	AJ532398
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	LAM	<i>Lepocinclis acus</i> var. <i>major</i>	UTEX LB 1316	apply for; cDNA	AF152104
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	LVO	<i>Lepocinclis ovum</i>	SAG 1244-8	apply for; cDNA	AF110419
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	LSGc38	<i>Lepocinclis spirogyroides</i>	SAG 1224-10b (clone 38)	apply for; cDNA	AY935694
	Euglenozoa/Euglenida/Euglenea	✓		✓			LSGc1_4	<i>Lepocinclis spirogyroides</i>	SAG 1224-10b (clones 1 & 4)	apply for; cDNA	-
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	LBU	<i>Lepocinclis steinii</i>	UTEX 523	apply for; cDNA	AF096993
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	LTR	<i>Lepocinclis tripteris</i>	UTEX 1311	apply for; cDNA	AY935696
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	MAA	<i>Monomorphina aenigmatica</i>	UTEX 1284	apply for; cDNA	AF190814
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	PYR	<i>Monomorphina pyrum</i>	UTEX 2354	apply for; cDNA	AF112874
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	MOU	<i>Monomorphina</i> sp.	MI 73	apply for; cDNA	DQ140130
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	POR	<i>Phacus orbicularis</i>	ACOI 996	apply for; cDNA	Triemer
	Euglenozoa/Euglenida/Euglenea	✓		✓	✓	✓	PSSc18	<i>Phacus similis</i>	SAG 58.81 (clones 18, 22,31)	apply for; cDNA	AJ532467
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	SAC	<i>Strombomonas acuminata</i>	SAG 1280-1 ACOI2467*	apply for; cDNA	DQ249883*
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	SBR	<i>Strombomonas borysteniensis</i>	NJ Stromb 10	apply for; cDNA	DQ140131
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	SNB	<i>Strombomonas napiformis</i>	NJ Strombo 2	apply for; cDNA	apply for; cDNA & gDNA
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	SOA	<i>Strombomonas ovalis</i>	NJ S115	apply for; cDNA	DQ140133
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	STB	<i>Strombomonas triquetra</i>	NJ S116	apply for; cDNA	DQ140153
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	TVS	<i>Trachelomonas conradi</i>	UTEX 1313	apply for; cDNA	AY015004
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	TEA	<i>Trachelomonas echinata</i>	SAG 1283-22	apply for; cDNA	AY015001
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	TLV	<i>Trachelomonas lefevrei</i>	SAG 1283-10	apply for; cDNA	DQ140136
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	TPL	<i>Trachelomonas planctonica</i>	Woongghi Trach 1 (WT1)	apply for; cDNA	DQ140138
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	TSM	<i>Trachelomonas similis</i>	SAG 1283-14	apply for; cDNA	DQ140142
	Euglenozoa/Euglenida/Euglenea/ “loricate” clade	✓		✓	✓	✓	TVI	<i>Trachelomonas volvocina</i>	UTEX 1327	apply for; cDNA	AJ532451
	Euglenozoa/Euglenida/Heteronematina	✓		✓	✓	✓	ENTO	<i>Entosiphon</i> sp.	Carolina	apply for; cDNA	AF220826

									Biological		
	Euglenozoa/Euglenida/Heteronematina	✓		✓			ENTSU	<i>Entosiphon sulcatum</i>	-	FJ807254; gDNA	-
	Euglenozoa/Euglenida/Heteronematina	✓		✓			PER	<i>Peranema trichophorum</i>	CCAP 1260/1b	apply for; cDNA	-
	Euglenozoa/Euglenida/Heteronematina	✓		✓	✓	✓	PERTR	<i>Peranema trichophorum</i>	CCAP 1260/1b	FJ807243; mRNA	AF386636
	Euglenozoa/Euglenida/Heteronematina	✓		✓			PLC	<i>Ploeotia costatum</i>	CCAP 1265/1	apply for; cDNA	-
	Euglenozoa/Euglenida/Heteronematina	✓		✓			PLO	<i>Ploeotia oblonga</i>	ATCC 50573	apply for; cDNA	-
	Euglenozoa/Euglenida/Heteronematina	✓		✓	✓	✓	PLOg	<i>Ploeotia oblonga</i>	ATCC 50573 (* = <i>P. costatum</i>)	apply for; gDNA	AF525486*
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	BST	<i>Bodo saltans</i>	CCAP 1907/2	FJ807252; gDNA	AY490233
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	DMT	<i>Dimastigella trypaniformis</i>	ATCC 50263	FJ807253; gDNA	AY028447
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida		✓				NEOSA	<i>Neobodo saliens</i> (syn. <i>Bodo saliens</i>)	ATCC 50358	FJ807267; mRNA	-
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	RBS	<i>Rhynchobodo</i> sp.	ATCC 50359	FJ807264; gDNA	U67183
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	RMN	<i>Rhynchomonas nasuta</i>	ATCC 50292	FJ807253; mRNA	AY425023
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida		✓				TRYBO	<i>Trypanoplasma borreli</i>	Tt-JH	FJ807268; mRNA	-
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	BLACU	<i>Blastocrithidia culicis</i>	ATCC 30268	FJ807251; gDNA	U05679
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	HERMU	<i>Herpetomonas muscarum</i>	ATCC 30260	FJ807255; gDNA	L18872
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	HERPE	<i>Herpetomonas pessoai</i>	ATCC 30252	FJ807256; gDNA	U01016
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	LEIBR	<i>Leishmania braziliensis</i>	MHOM/BR/75/M2904	XM_001563727; mRNA	GQ332355
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	LEIIN	<i>Leishmania infantum</i>	JPCM5	XM_001464627; mRNA	AM502245
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	LEIMA	<i>Leishmania major</i>	Friedlin	XM_001682206; mRNA	X53915
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	LEITA	<i>Leishmania tarentolae</i>	Parrot-II	apply for; gDNA	M84225
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓			LEITR	<i>Leishmania tarentolae</i>	UC	FJ807257; gDNA	-
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	LEPBI	<i>Leptomonas bifurcata</i>	- ; 53CR*	FJ807258; gDNA	DQ910925
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	LEPCO	<i>Leptomonas costaricensis</i>	- ; 15EC*	FJ807259; gDNA	DQ383648
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	LEPPO	<i>Leptomonas podlipaevi</i>	- ; clone 5-10-2*	FJ807260; gDNA	DQ383649
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	TRYAV	<i>Trypanosoma avium</i>	- ; A1412*	FJ807265; gDNA	U39578
	Euglenozoa/Kinetoplastida/ Metakinoplastina/Eubodonida	✓		✓	✓	✓	TRYBR	<i>Trypanosoma brucei</i>	TREU927	XM_817372; mRNA	M12676

	Euglenozoa/Kinetoplastida/ Metakinetoplastina/Eubodonida	√		√	√	√	TRYBE	<i>Trypanosoma brucei</i> <i>equiperdum</i>	STIB818	FJ807266; gDNA	AJ223564
	Euglenozoa/Kinetoplastida/ Metakinetoplastina/Eubodonida	√		√	√	√	TRYCR	<i>Trypanosoma cruzi</i>	CL Brener	XM_814346; mRNA	AF359495
	Euglenozoa/Kinetoplastida/ Metakinetoplastina/Eubodonida	√		√	√	√	TRYRA	<i>Trypanosoma rangeli</i>	San Agustin clone 1	AY785727; gDNA	AJ012417
	Euglenozoa/Kinetoplastida/ Prokinetoplastina	√		√	√	√	PERAM	<i>Perkinsiella amoebae</i>	AMOP1	FJ807262; gDNA	AY163355
	Euglenozoa/ <i>Incertae sedis</i> (Ma, 2005; Yubuki <i>et al.</i> , 2009)		√				-	<i>Petalomonas cantuscygni</i>	CCAP 1259/1	FJ807263; mRNA	-
	Fornicata/Eopharyngia/Diplomonadida	√		√	√	√; R	HEXIN	<i>Hexamita inflata</i>	AZ-4	U37081; U93087; gDNA	HXM16RR
	Fornicata/Eopharyngia/Diplomonadida	√		√	√	√	SPIVO	<i>Spironucleus vortens</i>	ATCC 50386	U94406; gDNA	SVU93085
	Heterolobosea/Acrasidae	√		√	√	√	ACRRO	<i>Acrasis rosea</i>	T-235	AF190771; gDNA	AF011458
	Heterolobosea/Gruberellidae	√		√	√	√	STALI	<i>Stachyamoeba lipophora</i>	-	FJ807246; mRNA	SLL00000112
	Heterolobosea/Vahlkampfiidae	√		√	√	√	NAEAN	<i>Naegleria andersoni</i>	A2	AF058283; gDNA	U80057
	Heterolobosea/Vahlkampfiidae	√		√	√	√	NAEGR	<i>Naegleria gruberi</i>	NEG-M	DQ295229; gDNA	AB298288
	Heterolobosea/Vahlkampfiidae	√		√	√	√	SAWMA	<i>Sawyeria marylandensis</i>	ATCC 50653	apply for; mRNA	AF439351
	Jakobida/Andalucia	√		√	√	√	ANDIN	<i>Andalucia incarcerationata</i>	MB1, Canada	EU334884; gDNA	EU334887
	Jakobida/Andalucia	√		√	√	√	ANDGO	<i>Andalucia godoyi</i>	AND28	EU334877; gDNA	AY965870
	Jakobida/Histionidae	√		√	√	√	HISAR	<i>Histiona aroides</i>	ATCC 50634	FJ807241; mRNA	HAL00000355
	Jakobida/Histionidae	√		√	√	√	RECAM	<i>Reclinomonas americana</i>	ATCC 50283	FJ807239; mRNA	AY117417
	Jakobida/Jakoba	√		√	√	√	JAKBA	<i>Jakoba bahamensis</i>	-	JBL00000183; mRNA	JBL00000661, JBL00000818, JBL00000193
	Jakobida/Jakoba	√		√	√	√	JAKLI	<i>Jakoba libera</i>	ATCC 50422	FJ807244; mRNA	AY117418
	Jakobida/ <i>Incertae sedis</i>	√		√	√	√	SECEC	<i>Seculamonas ecuadoriensis</i>	ATCC 50688	FJ807245; mRNA	DQ190541
	<i>Malawimonas</i>	√		√	√	√	MALCA	<i>Malawimonas californiana</i>	ATCC 50740	FJ807242; mRNA	MCL00000562
	<i>Malawimonas</i>	√		√	√	√	MALJA	<i>Malawimonas jakobiformis</i>	ATCC 50310	DQ295227; gDNA	AY117420
	Parabasalia	√		√			HISME	<i>Histomonas meleagridis</i>	-	FM200063; mRNA	-
	Parabasalia	√		√	√	√	TRIVA	<i>Trichomonas vaginalis</i>	clone G3	XM_001328923; mRNA	AY338476
	Preaxostyla/Oxymonadida	√		√	√	√	DINEN	<i>Dinenympha exilis</i>	A-5	AB007028; mRNA	AB092924

	Preaxostyla/Oxymonadida	√		√	√	√	MONSP	<i>Monocercomonoides</i> sp.	clone PA	EF474125; mRNA	AY831435
	Preaxostyla/Oxymonadida	√		√	√	√	STRST	<i>Streblomastix strix</i>	clone 1	AY188861; gDNA	AY188886
	Preaxostyla/ <i>Trimastix</i>	√		√	√	√	TRIPY	<i>Trimastix pyriformis</i>	ATCC 50562	DQ295235; gDNA	AF244903
Opisthokonta	Choanomonada	√		√	√		MONOV	<i>Monosiga ovata</i>	-	AY582824; mRNA	AF271999
	Fungi/Ascomycota	√		√	√		NEUCR	<i>Neurospora crassa</i>	OR74A	XM_959775; mRNA	X04971
	Fungi/Basidiomycota	√		√	√		CRYNE	<i>Cryptococcus neoformans</i>	JEC21	XM_568462; mRNA	L05428
	Fungi/Chytridiomycetes/ Monoblepharidales	√		√	√		HARSP	<i>Harpochytrium</i> sp.	JEL94	AF450113; gDNA	AF164331 AF164332
	Fungi/Zygomycota/Dimargaritales	√		√	√		ABSG	<i>Absidia glauca</i>	CBS 101.48	X54730; gDNA	AF157118
	Mesomycetozoa	√		√	√		CAPOW	<i>Capsaspora owczaraki</i>	-	DQ403163; mRNA	AF349564
	Metazoa/Animalia	√		√	√		ORNAN	<i>Ornithorhynchus anatinus</i>	Glennie (liver, lung, brain, spleen), ♀	XM_001507841; mRNA	AJ311679

Table 2.3. Summary of alignments. Alignment short names are abbreviations of the gene(s), number of sequences, and type of positions included in the alignment as detailed in separate columns of the table. Genes are EF1 α =1a or EF1 α with third codon positions removed=1a3rd, small subunit ribosomal DNA=18S and the concatenated alignments utilizing both EF1 α and 18S are either 1a18S or 1a3rd18S. OTU#=operational taxonomic units number used in each alignment where “101” includes all Excavates (including repeated OTUs where sequences exhibited >5% sequence identity), “92” excludes duplicate Excavate OTUs (*Leishmania tarentolae*, *Euglena gracilis* (X16890), *E. gracilis* (FJ807238), *Peranema trichophorum* (applied for), *Entosiphon sulcatum*, *Lepocinclis spirogyroides* (clones 1 & 4), *Ploeotia costatum* (18S from *P. costatum* used in concatenation with EF1 α gDNA sequence from *P. oblonga*), *Ploeotia oblonga* (cDNA)) and any organisms that did not have an 18S (*Histomonas meleagridis*), and “84” excludes all OTUs excluded from alignment “92” as well as all unikont outgroup organisms. Alignments with the 3rd codon nucleotide positions removed in the EF1 α sequences are denoted with an “X”. Number of base pairs (bp) in each alignment and which OTU was forced as the root outgroup in each analysis using that alignment are also described. Excluded blocks in the brackets are based on the Mo nucleotide alignment of 194 OTUs (Fig. 1.1) for EF1 α . Included OTUs compared to Figs. 1.1, 2.1 and 2.2 are marked in Table 2.2.

Short Name	Gene(s)	OTU#	Excluded Blocks	3 rd positions removed	bp	Outgroup
1a-101	EF1 α	101	[1-69], [433-456], [550-573], [634-636], [643-717], [721-738], [805-816], [829-834], [844-900], [931-933], [991-993], [1093-1119], [1210-1212], [1252-1260], [1264-1266], [1273-1275], [1288-1299], [1333-1335], [1357-1389], [1453-1461], [1480-1488], [1504-1515], [1588-1560], [1627-1803]		1227	<i>Acanthamoeba culbertsoni</i> (Amoebozoa)
1a3 rd -101	EF1 α	101		X	818	<i>Acanthamoeba culbertsoni</i> (Amoebozoa)
1a-92	EF1 α	92			1227	<i>Acanthamoeba culbertsoni</i> (Amoebozoa)
1a3 rd -92	EF1 α	92		X	818	<i>Acanthamoeba culbertsoni</i> (Amoebozoa)
1a-84	EF1 α	84			1227	<i>Hexamita inflata</i> (Excavata/Fornicata)
1a3 rd -84	EF1 α	84		X	818	<i>Hexamita inflata</i> (Excavata/Fornicata)
18S-92	18S	92	n/a	n/a	1008	<i>Acanthamoeba culbertsoni</i> (Amoebozoa)
18S-84	18S	84	n/a	n/a	1048	<i>Hexamita inflata</i> (Excavata/Fornicata)
1a18S-92	EF1 α & 18S	92	same as 1a excluded blocks above for EF1 α portion of concatenation		2238	<i>Acanthamoeba culbertsoni</i> (Amoebozoa)
1a3 rd 18S-92	EF1 α & 18S	92		X	1828	<i>Acanthamoeba culbertsoni</i> (Amoebozoa)
1a18S-84	EF1 α & 18S	84			2275	<i>Hexamita inflata</i> (Excavata/Fornicata)
1a3 rd 18S-84	EF1 α & 18S	84		X	1866	<i>Hexamita inflata</i> (Excavata/Fornicata)

Table 2.4. Summary of tree statistics from empirical data and results from ML and MLE analyses. Analysis details and parameters are listed under the analysis type and number of trees used to calculate support is stated (# trees).

Analysis Type & Parameters	Alignment Input	# trees
<u>RAxML (ML)</u> autodetect bootstrap replicates 1a & concatenated = GTR+I+ Γ 18S= GTR+ Γ	1a-101	300
	1a3 rd -101	400
	1a-92	300
	1a3 rd -92	400
	1a-84	300
	1a3 rd -84	450
	18S-92	300
	18S-84	250
	1a18S-92	200
	1a3 rd 18S-92	200
	1a18S-84	200
	1a3 rd 18S-84	200
<u>MrBayes (MLE)</u> McMC parameters: nchains=4; Temp=0.1; Swapfreq=1; Nswaps=1 GTR+I+ Γ , except 18S alignments (GTR+ Γ); 4 gamma categories; estimated empirical base frequency 1a = 30 million generations with 15 million burn-in; 4 runs each with trees after burn-in combined to make one .con tree 18S & concatenated = 4 million generations with 2 million burn-in; two analyses of 2 runs each to check for topology convergence	1a-101	60,000
	1a3 rd -101	60,000
	1a-92	60,000
	1a3 rd -92	60,000
	1a-84	60,000
	1a3 rd -84	60,000
	18S-92	20,000
	18S-84	20,000
	1a18S-92	20,000
	1a3 rd 18S-92	20,000
	1a18S-84	20,000
	1a3 rd 18S-84	20,000

Table 2.5. Summary of recovered clades from all phylogenetic analyses with Euglenozoa retaining EF1 α . A ++ denotes a strongly supported clade (both analysis types recover it), + denotes a supported clade (one analysis type recovers it) and a blank spot denotes it was not ever recovered or it was not possible to recover it due to reduced OTUs studied in the given alignment. Analysis types were RAXML (ML; “best tree” used with rapid bootstrap support) and MrBayes (MLE; consensus tree). Any clades that are represented by two or more OTUs are included as subcategories. The following subcategories are reported because they have been recovered as clades in previous studies: unikonts, bikonts, CMM (Choanomonada + Mesomycetozoa (specifically Ichthyosporea) + Metazoa), “Group 1” Excavates, “Group 2” Excavates, “Group 3” Excavates, Discicristates, Eutreptiales, and the “loricate” clade. Alignment numbers correspond to number of OTUs used in each alignment where “101” includes all Excavates (including repeated species where sequences exhibited <95% sequence identity), “92” excludes duplicate Excavate OTUs (*Leishmania tarentolae*, *Euglena gracilis* (X16890), *E. gracilis* (FJ807238), *Peranema trichophorum* (applied for), *Entosiphon sulcatum*, *Lepocinclis spirogyroides* (clones 1 & 4), *Ploeotia costatum* (18S from *P. costatum* used in concatenation with EF1 α gDNA sequence from *P. oblonga*), *Ploeotia oblonga* (cDNA)) and any organisms that did not have an 18S (*Histomonas meleagridis*), and “84” excludes all OTUs excluded from alignment “92” as well as all unikont outgroup organisms. Genes are EF1 α =1a or EF1 α with third codon positions removed=1a3rd, small subunit ribosomal DNA=18S and the concatenated alignments utilizing both EF1 α and 18S are either 1a18S or 1a3rd18S. If Euglenozoa was recovered as a monophyletic clade, and if the group sister to Euglenozoa was also a recognized, recovered monophyletic clade in both analysis types, then an asterisk (*) was put next to the sister clade in the same column of that alignment.

Clade	Alignment Information: Number of Taxa and Genes Used in Alignment											
	101	92			84			101	92			84
	1a	1a	18S	1a18S	1a	18S	1a18S	1a3rd	1a3rd	1a3rd18S	1a3rd	1a3rd18S
Ingroup												
****Euglenozoa			++	++		++	++	++	++	++	++	++
*****Euglenida								++	++		++	+
*****Euglenea		++	++	++	++	++	++		++	++	++	++
*****Colacium			++	++		++	++			++		++
*****Cryptoglana	++	++	++	++	++	++	++			++		++
*****Euglena												
*****Eutreptiales			++			++				++		+
*****Lepocinclis			++	++		++	++			++		++
*****Monomorphina				++		++	++					
*****Phacus				++			++					++
*****“loricate” clade												
*****Strombomonas			++	++		++	++			++		++
*****Trachelomonas												
*****Heteronematina									+			

***** <i>Entosiphon</i>	++							++				
***** <i>Peranema</i>	++							++				
***** <i>Ploeotia</i>												
*****Kinetoplastida	++	++	++	++	++	++	++	++	++	++	++	++
*****Metakinetoplastina			++	++		++	++					
*****Neobodonida			++			++						
*****Trypanosomatida	++	++		++	++		++	++	++	++	++	++
***** <i>Herpetomonas</i>	++	++	++	++	++	++	++	++	++	++	++	++
***** <i>Leishmania</i>		++	+	++	++	+	++			++		+
***** <i>Leptomonas</i>										+		
***** <i>Trypanosoma</i>	++	++	++	++	++	++	++			++		++
Outgroup												
* “bikonts”/Excavata	++	++	++	++				++	+	++		
** “Group 1”	++	++		++	++	++	++	++	++	++	+	++
***Fornicata/Eopharyngia/ Diplomonadida	++	++	++	++	++	++	++	++	++	++	+	++
***Parabasalia	++							++				
** “Group 2”						++					+	+
*** <i>Malawimonas</i>	++	++	++	++	++	++	++			++		++
***Preaxostyla	++	++	+	++	++	++	++	++	++	++	++	++
****Oxymonadida	++	++	+	++	++	++	++	++	++	++	++	++
** “Group 3”; “Discoba”	++	++		++			++	++	++	++	++	++
***Jakobida				++			++					
**** <i>Andalucia</i>			++	++	+	++*	++			++		++
**** Histonidae			++	++		++	++			++		++
**** <i>Jakoba</i>												
***Discicristates				++			++			++		++
****Heterolobosea			++	++*		++	++*			++*	++	++*
*****Vahlkampfiidae	++	++		++	++		++	++	++		++	
***** <i>Naegleria</i>	++	++	++	++	++	++	++	++	++	++	++	++
* “unikonts”	++	++	++	++				++	+	++		
**Opisthokonta	++	++	++	++				++	++	++		
***CMM	++	++		++				++	++	++		
***Fungi	++	++		++				+		++		
Total recovered clades	21	20	24	32	15	24	27	20	18	29	15	26

Figure 2.1. The 18S-92 interleaved alignment. The OTUs are grouped according to supergroup, then to “group” that were related in previous literature. The OTUs are given an abbreviated short name (see Table 2.2) which is consistent throughout the pages of the alignment. The column positions, from 1-1048, are given in ten base increments at the top of each page. The number of residues from the original sequence is given in brackets to the right of each line corresponding to its respective OTU.

	10	20	30	40	50	60	70	80	90	100	110	120	
ACACU	G-ATTAAGCC	ATGCATGTCT	AGACTGCGGA	TGGCTCATT	AATCAGTTAT	AATTCTAGAG	CTAATACTTC	ATTCAAATTT	CTGCCCTATC	AGATGGTAGG	ATAGAGGCCCT	ACCATGGTCG	[119]
CAPOV	G-ATTAAGCC	ATGCATGTCT	AAACTGCGAA	TGGCTCATT	CATCAGTTAT	AATTCTAGAG	CTAATACTATC	ATTCAAATTT	CTGCCCTATC	AGATGGTAAG	GTATTGGCCT	ACCATGGTTA	[119]
MONOV	G-AtTAagCc	AtGcAtGtCt	AaAcTgCgAa	TgGcTcAtTa	AaTcAgTtAt	AaTtCtAgAg	CtAaTaCaTc	AtTcAaAtTt	CtGcCcTaTc	AgAtGgTaAg	GtAtTgGcCt	AcCaTgGtCg	[119]
ORNAN	G-ATTAAGCC	ATGCATGTCT	AAACTGCGAA	TGGCTCATT	AATCAGTTAT	AATTCTAGAG	CTAATAACCC	ATTGGAACGT	CTGCCCTATC	AGATGGTACT	TGCCGTGCCCT	ACCATGGTGA	[119]
NEUCR	G-ATTAAGCC	ATGCATGTCT	AAACTGCGAA	TGGCTCATT	AATCAGTTA-	AATTCTAGAG	CTAATACTTC	ATTCAAATTT	CTGCCCTATC	AGACGGCTGG	GTCTTGGCCA	GCCATGGTGA	[118]
CRYNE	G-ATTAAGCC	ATGCATGTCT	AAACTGCGAA	TGGCTCATT	AATCAGTTA-	AATTCTAGAG	CTAATACTTC	ATTCAAATTT	CTGCCCTATC	AGATGGTAGG	ATAGAGGCCCT	ACCATGGTAT	[118]
HARSP	-----	-----	-AACTGCGAA	TGGCTCATT	AATCAGTTA-	AATTCTAGAG	CTAATACTTC	ATTCAAATTT	CTGCCCTATC	AGATGGTACG	GTATTGGCCT	ACCATGGTTT	[98]
ABSGL	G-ATTAAGCC	ATGCATGTCT	AAACTGCGAA	TGGCTCATT	AATCAGTTA-	AATTCTAGAG	CTAATACTAC	ATTGATTTT	CTGCCCTATC	AGATTGTAAG	GTAGTGGCTT	ACAATGGCGA	[118]
HEXIN	GGATTAAGCC	ATGCATGCC-	AAACTGCGAA	CGGCTCATT	AATCAGTTGA	AACCTCTCTTA	CTAA-TCGCA	CGATGACTT	CTGACGCATC	AGTAGGAGAA	GTAAAGGTTA	TCCTATGAGT	[118]
TRIPY	G-ATTAAGCC	ATGCATGTCT	AAACTGCGAA	TGGCTCATT	TATCAGTAAT	AATTCTAGAG	CTAATACTATC	ATTCAAATTT	CTGCCCTATC	AGATGGTAGG	ATAGAGGCCCT	ACCATGGTTT	[119]
MALJA	G-ATTAAGCC	ATGCAAGCCC	-ATCTGCGAA	TGGCTCATT	AATCAGTTAT	AATTCTAGGG	CTAATACTATC	ATTCAAATTT	CTGCCCTATC	AGACGGTARG	ATAGAGGCTT	ACCGTGGTCG	[118]
MALCA	C-ATTAAGCC	ATGCAAGCCC	-ATCTGCGAA	TGGCTCATT	AATCAGTTAT	AATTCTAGGG	CTAATACTATC	ATTCAAATTT	CTGCCCTATC	AGACGGTAAG	ATAGAGGCTT	ACCGTGGTCG	[118]
DINEX	G-ATTAAGCC	ATGCAAGTCT	ACACTGCGAA	AAGCTCCTTA	AAATAGTTAT	AATTCTAGAG	CTAATACTTC	ATTCAAATTT	CTAACCTATC	AGACGTTAAG	GTAGTGTCTT	AGAGTAGTCG	[119]
MONSP	T-TTGAATAG	AAATATAATT	TTAGAGCTAA	TACATGCGAG	AATTCATTTT	AAAGATCCGA	CTAAGGGGTC	ATTGAAATTT	CTGACCTATC	AGATGGTAGG	ATAGAGGCCCT	ACCATGGTAT	[119]
STRST	G-ATTAAGCC	ATGCATGTGT	CAACTGCGAA	TGGCTCATT	TATCAGTTAT	AATTCTAGAG	CTAATACTATC	ATTGAAATTT	CTGCCCTATC	AGACTTAAGG	GTAGTTTCCCT	TAGTGGTAT	[119]
TRIVA	TATTTAAGCC	ATGCAAGTGT	TAAGTGCAG	TAGCTCATT	ATACGCTCAG	AACCTCTGGT	CTAATACGCC	ATTGCACTGA	GTGACCTATC	AGTACTTAGG	GTCTTTACCT	AGGTAGGCTA	[120]
SPIVO	G-ACTAAGCC	ATGCAAGCC-	ACTC-GCAAA	CGGCTCATT	TACGCGTCTT	AACCTCTCTT	CCAAATCCGC	ATTG-TGATT	CGCAGCATC	AGTAGGTAGC	GTAAAGTGTCT	TCTTAGGGAT	[116]
ANDIN	-----	-----TGCT	CAACTGCGAA	TGGCTCATT	AATCAGTGAT	AACCTGTAGAG	CTAATACTATC	ATTGGAATTT	CTGCCCTATC	AGATGGTAGT	GTGCTGGACT	ACCATGGCGA	[105]
ANDGO	G-ATTAAGCC	ATGCATGTCT	AATCTGCGAA	CGGCTCATT	AACAAGCTAT	AACGTGTAGAG	CTAATACTATC	ATTCAAATTT	CTTCCCTATC	AGACGGTAGT	GTAGTGGACT	ACCGTGGCGG	[119]
SECEC	AGATTAAGCC	ATGCATGTCT	AATCTGCGAA	CGGCTCATT	CAACAGTCAT	AACAGCAGAG	CTAACACACC	AATCAGGCT	CTGCCCTATC	AGTCCGTACG	GTATTGGCCT	ACCGAGGCCG	[120]
RECAM	GGATTAAGCC	ATGCATGTCT	AAACTGCGAA	CGGCTCATT	CAACAGTTAT	AACACCAGAG	CTAATAACACC	AATCGAGTTT	CTGCCCTATC	AGTCCGTACG	GTATTGGCGT	ACCGAGGCCG	[120]
HISAR	GGATTAAGCC	ATGCATGTCT	AAACTGCGAA	CGGCTCATTT	CAACAGTCAT	AACACCAGAG	CTAATAACACC	AATCAAGTTT	CTGCCCTATC	AGTCCGTTGG	ATATTGCGCA	ACCGAGGCCG	[120]
JAKLI	AGATTAAGCC	ATGCATGTCT	AAACTGCGAA	CGGCTCATT	CAACAGTTAT	AACCTCCAGAG	CTAATAACACC	AATCAAGTTT	CTGCCCTATC	AGTCCGTAGG	GTATTGGCCT	ACCGAGGCCG	[120]
JAKBA	GGATTAAGCC	ATGCATGTGT	CAACTGCGAA	CGGCTCATT	CAACAGTCAT	AACCTCCAGAG	CTAATAACACC	AATCGAGTCT	CTGCCCTATC	AGACGGTAGT	GTATTGGACT	ACCGTGGCTT	[120]
STALI	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
NAEGR	AGCCTAAGCC	ATGCAAAATGT	AGTCTGTGGA	AGGCTCATT	TAACAGTTAT	AACCTGCAAGC	AT-ATACGGC	AGAGGAGTTT	CTTACCTATC	AGTTGTTTGT	TTAAAGGACA	AACAGGCCCT	[119]
NAEAN	AGCCTAAGCC	ATGCAAAATGT	AGTCTGTGGA	AGGCTCATT	TAACAGTTAT	AACCTGCAAGC	ATAATACGGC	AGAGGAGTTT	CTTACCTATC	AGTTGTTTGT	TTAAAGGACA	AACAGGCCCT	[120]
ACRRO	AGATTAAGCC	ATGCATGTCT	AATCTGTGAA	CGGCTCATT	TAACAGTTAT	AAATGCGAGC	ATAATACGGA	AGACGAATTT	CTTATCTATC	AGTTGTCTCA	ATAGGGGTGA	GACACGGCGA	[120]
SAWMA	G-ATTAAGCC	ATGTAATATGT	AAACTGTGCA	AGGCTCAATA	AACGATTTAT	AAAAGCATCG	ATAATACAGT	CGATGATTTT	ATCATCTATC	AATAAGTGT	--TAATTTAC	ACTCAGACAA	[117]
BST	GGACTAAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
RBS	GGACTAAGCC	ATGCACGCC	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACG-CAAG	CTAATACTTC	GA-CGAACAA	CTGCCCTAT-	-----CGT	GTAGTGGACT	GCCATGGCTT	[110]
RMN	GGACTAAGCC	ATGCATGCCT	CGTCTGCGCA	TGGCTCATT	CATCAGAGCG	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
DMT	GGACTAAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCG	AAACGCCGAGG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
PERAM	-----	-----	-----	-----	-----	AAACGCCAAG	CTAATACTTT	GA-CGAGTCT	CTGCCCTATC	AGATGGTACT	GTATTGGACT	ACCATGGCGA	[69]
LEIMA	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
LEIIN	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAAC-A	CTGCCCTATC	AGATGCC-CT	GTAAATGGACT	GCCAT-GCTT	[116]
LEIBR	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
LEITA	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
LEPCO	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
LEPBI	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	A--TGCCGT	GTAGTGGACT	GCCAGGGCGT	[117]
LEPPO	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
HERMU	GGACTTAGCC	ATGCATGCCT	CGTCTGCGCA	TGGCTCATT	CATCAGAGCT	AAAGG-CAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[118]
HERPE	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACG-CAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[118]
BLACU	GGACTAAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCTAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACG	GCCATGGCGT	[119]
TRYBR	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTATC	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
TRYBE	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTATC	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
TRYAV	GGGCTTAGC-	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[118]
TRYCR	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATTTG	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
TRYRA	GGACTTAGCC	ATGCATGCCT	CATCTGCGCA	TGGCTCATT	CATCAGAGCT	AAACGCCAAG	CTAATACTTT	GA-CGAACAA	CTGCCCTATC	AGATGGCCGT	GTAGTGGACT	GCCATGGCGT	[119]
ENTO	GGACTAAGCC	ATGCATATGC	GTCTGTGGA	CGGCTCATT	TTGCGATTGC	AACCC-ACCC	TTAATACAGG	ATTTCG-AGT	CGACCCATATC	AGAAGGCAAG	GTCTTGGCCT	GCCTTGACC	[118]
PLog	G-ATTAAGCC	ATGCAAACTT	CGTCCGTGAA	CGGCTCATT	AATCAGGATG	AACCTGTATCG	GACGAAATTT	GTAGTGTATC	CTGCCCTATC	AGATGACTGT	GCAATGGAAC	ACAATGGTAT	[119]
PERTR	GGACTAAGCC	ATGCAAGTGT	AGTCTGCGAA	TGGCTCATT	CAACAGCAAC	AACCGGCTAG	TTAATACGTC	ATAGGAACTT	CTGCCCTATC	AGATGGTAGG	GTAGTGGCTC	ACCATGGCAT	[120]
ESL	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGAGG	CTAATACTATC	GCTCGAACTT	CTGACTATATC	AGATTGTGGT	GTAGTGGACC	ACAATGGCTC	[120]
EXS	GGACTAAGCC	ATGCATATCT	CATCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACTATC	GACGGAACCT	CTGACCTATC	AGACTGTGGC	GCAAGGGCCC	ACAAGTGGCG	[120]
SAC	-----	-----	-----TTA	CATCAGCAGT	AACCTAGTGG	CTAATACTATC	ACTCGAACTT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAAGTGGCA	[83]	
ECL	TGCAGGTGTT	GTGCCGTGATT	CGGCTCGGAT	TTGATCGGAA	GAGCAGCGTT	AGTCCGCCCC	GCACCACATC	ACGCGAACTT	CTGACCTATC	AGACTGTGGT	GTAGCGGACA	CAAGTGGCAG	[120]
TSM	GGACTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACTATC	ACACGAACTT	CTGACCTATC	AGACTGTGGT	GTAGCGGACC	ACGATGGCTA	[120]

	10	20	30	40	50	60	70	80	90	100	110	120	
POR	-----	ATGCAAGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	GCACGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCTA	[110]
TVI	-----	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACATC	ACACGAACCT	CTGACCTATC	AGACTGTGGT	GTAGCGGACC	ACAGTGGCTT	[110]
MAA	GGGCTAAGCC	ATGCACGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	ACACCAACTT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCGG	[120]
PSSc18	-----	ATGCATGTCT	CATCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	ACACGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCAG	[110]
EDS	GGGCTAAGCC	ATGCAAGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACATC	GCACGAACCT	CTGACCTATC	AGACTGTGGT	GTCTGGGACC	ACAGTGGCCT	[120]
MOU	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	ACAAGAACTT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCAG	[120]
PYR	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	GCAAGAACTT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCCG	[120]
LVO	GGACTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGG	AACCTAGTGG	CTAATACATC	GCACGAGCTT	CTGACCTATC	AGATGGTGGT	GTAGAGCCC-	ACCATGGCGG	[119]
TEA	GGATTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACATC	ACATGAACCT	CTGACCTATC	AGACTGTGGT	GCAGTGGACC	ACAGTGGCCT	[120]
EGR	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	ACACGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCGG	[120]
EPE	GGACTAAGCC	ATGCATGTCT	CATCTGTGAA	TGGCTCCTTA	CATCAGCAAC	AACCTGGTGG	CTAATACTTG	GAGCGAACTT	CTGACCTATC	AGATTGTGGT	GTCTGGGACT	ACAATGGCTT	[120]
TLV	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	ACACGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCAG	[120]
TPL	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	ACACGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCAG	[120]
LBU	-----	ATGCATGTCT	GGTCTGTGAA	TGGCTCCTTA	CATCAGCAGG	AACCTAGTGG	CTAATACATC	ACACGAACCT	CTGACCTATC	AGATGGTGGT	GTCTAGGCC-	ACCATGGCAG	[109]
DAA	GGGCTAAGCC	ATGCATGTCT	CATCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTAGTGG	CTAATACACC	ACATGAACCT	CTGACCTATC	AGATTGTGGT	GTAGTGGACC	ACAATGGCCT	[120]
LAM	GGACTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	ACACGAGCTT	CTGACCTATC	AGATGGTGGT	GTATTGGACC	ACCATGGCAG	[120]
LTR	-----	-----	-----	-----TTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	ACACGAACCT	CTGATCTATC	AGATGGTGGT	GTAGTGGACC	ACCATGGCAG	[83]
EUGLO	-----	ATGCACGTCT	CATCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACATC	ACAAGAGCTT	CTGACCTATC	AGACTGTGGT	GTATTGGACC	ACAGTGGCCT	[110]
EVI	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACATC	ACATGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCTA	[120]
LSGc38	-----	-----	-----	-----TTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	ACAAGAACTT	CTGACCTATC	AGATGGTGGT	GTAGTGGACC	ACCATGGCGG	[83]
CPI	-----	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	ACAAGAACTT	CTGCCCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCAA	[110]
CSK	-----	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	ACAAGAACTT	CTGCCCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCAG	[110]
TET	-----	ATGCAAGCCT	CATCTGTGAA	TGGCTCCTTA	CATCAGCAAT	AACCTGGTGG	TTAATACACC	AGACGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACCGTGGCTA	[110]
EGN	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACGTC	ACAGGAACCT	CTGACCTATC	AGATTGTGGT	GTAGTGGACC	ACAATGGCTT	[120]
TVS	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACATC	ACACGAACCT	CTGACCTATC	AGATTGTGGT	GTAGTGGACC	ACAATGGCAG	[120]
SNB	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	ACTCGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCCG	[120]
EHS	GGGCTAAGCC	ATGCACGTCT	CATCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACATC	GCAAGAGCTT	CTGACCTATC	AGACTGTGGT	GTATTGGACC	ACAGTGGCCT	[120]
GRC	-----	ATGCACGTCT	CATCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTTGTGG	CTAATACATC	GCAAGAGCTT	CTGACCTATC	AGACTGTGGT	GTATTGGACC	ACAGTGGCCT	[110]
STB	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	ACTCGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCCA	[120]
MUC	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACACC	ACAGGAACCT	CTGACCTATN	AGATTGTGGT	GTAGTGGACC	ACAATGGCAT	[120]
CVL	-----	-----	-----	-----TTA	CATCAGCAGT	AACCTGGTGG	CTAATACACC	ACAGGAACCT	CTGACCTATC	AGATTGTGGT	GTAGTGGACC	ACGATGGCAT	[83]
SOA	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	ACTCGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCAG	[120]
SBR	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTAGTGG	CTAATACATC	ACTCGAACCT	CTGACCTATC	AGACTGTGGT	GTAGTGGACC	ACAGTGGCAA	[120]
CEL	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACTCC	ACAGGAACCT	CTGACCTATC	AGATTGTGGT	GTAGTGGACC	ACAATGGCAT	[120]
CUA	GGGCTAAGCC	ATGCATGTCT	CGTCTGTGAA	TGGCTCCTTA	CATCAGCAGT	AACCTGGTGG	CTAATACACC	ACAGGAACCT	CTGACCTATC	AGATTGTGGT	GTAGTGGACC	ACAATGGCAT	[120]

	130	140	150	160	170	180	190	200	210	220	230	240		
ACACU	TAACGGGTAA	CGGAGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATGGCTACCA	CTTCTAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACAATAAAAC	[239]	
CAPOW	CAACGGGTAA	CGGAGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ATGGCTACCA	CATCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACAATAAAAC	[239]	
MONOV	CaAcGgGtAa	CgGgGaAtTa	GgGtTcGgTt	CcGgAgAgGg	GgCcTgAgAa	AtGgCcAcCa	CaTcCaAgGa	AgGcAgCaGg	CgCgCaAaTt	AcCcAaTcCg	GgAgGtAaTg	ACaAaTaaAc	[239]	
ORNAN	CCACGGGTGA	CGGGGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CATCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACGAAAAAAC	[239]	
NEUCR	CAACGGGTAA	CGGAGGGTTA	GGGCTCGACC	CCGGAGAAGG	AGCCTGAGAA	ACGGCTACTA	CATCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACAATAAACT	[238]	
CRYNE	CAACGGGTAA	CGGGGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CATCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACAATAAAAC	[238]	
HARSP	TAACGGGTAA	CGGGGAATTA	GGGTTTCGTT	CCGGAGAGGG	-ACCTGAGAA	ATGGCTACCA	CATCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCTA	GGAGGTAGTG	ACAAGAAAAAC	[217]	
ABSLG	CCACGGGTAA	CGGGGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CATCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACAATAAAAC	[238]	
HEXIN	TCACGCGT - A	CGGAGAATTA	GGGTTTCGACT	CCGGAGAAATG	AGCATGAGAG	ACGGCTACCA	GTCTTAAGGG	AGGCAGCAGG	CGCGAAATT	GCCCAATGTA	CGAGGCAGTG	AGAAAAA - CT	[236]	
TRIPY	TGACGGGTAA	CGGAGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAAAG	ATGGCGACTA	CTTCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACAAAAAATC	[239]	
MALJA	TCACGGGTAG	CGGAGGATTA	GGGTCTGACT	CCGGAGAGGG	AGCTTTAGAA	ACGGCCGCCA	CTTCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATTCG	AGAGGTAGTG	AAGACAAAAC	[238]	
MALCA	TCACGGGTAG	CGGAGGATTA	GGGTCTGACT	CCGGAGAGGG	AGCTTTAGAA	ACGGCCGCCA	CTTCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATTCG	AGAGGTAGTG	AAGACAAAAC	[238]	
DINEX	TGACGGGTGA	CGGGGAATCA	GGGTTTCGATA	CCGGAGAGGG	CGCATGAGAA	ATGGCGACCA	TCACTAAAGT	GAGCAGCAGG	CGCGCAAATT	ACCCAATGCG	TGAGGTAGTG	ACAAAAAAC	[239]	
MONSP	TGACGGGTAA	CGGAGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ATGGCTACCA	CTTCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACAAAAACAA	[239]	
STRST	TGACGGGTAA	CGGAGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACAGCTCCCA	CTTCTAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGTG	ACAAAAAAC	[239]	
TRIVA	TCACGGGTAA	CGGGCGGTTA	CCGTGCGACT	CCGGAGAAGG	CGCCTGAGAG	ATAGCGACTA	TATCCACGGG	TAGCAGCAGG	CGCGAAACTT	TCCCACTCGC	GGAGGAGGTA	ATGACCA - TC	[239]	
SPIVO	TAACGCCT - A	CAGGACATTA	GGGTGCGACT	CTGGAGAGTG	AGCATGAGAG	ACGGCTCACA	GTCTTACGGA	AGGCAGCAGG	CGCGGAAATT	GCCCAATGTA	CGAGGCAGTG	AGAAGCATGG	[235]	
ANDIN	TAACGGGTAA	CGGAGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATGGCTACCA	CATCCAAGGA	AGGCAGCAGG	CGCGTAAATT	ACCCAATCCG	GGAGGTAGTG	ACCATAAATG	[225]	
ANDGO	TAACGGGTAA	CGGAGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTTAAAA	ATGGCTACCA	CAACTAAGGT	TGGCAGCAGG	CGCGCAACTT	ACTGACTCTA	GGAAGTAGTG	ACAATAAAAC	[239]	
SECEC	CGACGGGTAA	CGGGGAATTA	GGGTTTGATT	CCGGAGAGGG	AGCCTGAGAG	ATGGCTACCA	CATCCAAGGA	AGGCAGCAGG	CGCGCAAATT	GCCCAATCCG	GGAGGCAGTG	ACGAGAAAAAC	[240]	
RECAM	CAACGGGTAA	CGGGGAATTG	GGGTTTGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CATCTAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATGCG	CGAGGTAGCG	ACAAGAAAAAC	[240]	
HISAR	CCACGGGTAA	CGGGGGATTA	GGGTTTGATT	CCGGAGAGGG	GGCCTGAGAG	ATGGCCACCA	CATCTAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATGCG	CGAGGTAGCG	ACAAGAAATC	[240]	
JAKLI	CTACGGGTAA	CGGGGAATTA	GGGTTTGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CATCTAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCTA	GGAGGTAGTG	ACAAGAAACC	[240]	
JAKBA	CTACGGGTAA	CGGGGAATTA	GGGTTTGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CATCCAAGGA	AGGCAGCAGG	CGCGCAAATT	ACCCAATCCG	GGAGGTAGCG	ACAAGAAAAAC	[240]	
STALI	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]	
NAEGR	TGACGGGT - A	CGGGGAATCA	GTGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATCGCTACCA	CATCTAAGGA	CGGCAGCAGG	CGCGCAAATT	ACCCAATCTA	GGAGGTAGTG	ACAAGCTAGT	[238]	
NAEAN	TTACGGGT - A	CGGGGAATCA	GTGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATCGCTACCA	CATCTAAGGA	CGGCAGCAGG	CGCGCAAATT	ACCCAATCTA	GGAGGTAGTG	ACAAGATAGT	[239]	
ACRRO	TTACGGAT - T	CGGAGAATCA	GTGTTTGATT	CCGGAGAGGG	AGCCTTAAAA	ATGGCTACTA	CATCTAAGGA	CGGCAGCAGG	CGCGAAAATT	GCCCAATCTA	GGAGGCAGTG	ACAAGAAAAAC	[239]	
SAWMA	CCACAGAT - A	CGGATAATCA	GTGTTTGATT	CCGGAGAAAG	AGCCTTAAAA	ATGGCTACTA	GTACCAAGGT	TTGCAACAGG	CGCGAAAATT	ATCCAATGCG	CGAGATAGTG	AAAAGAAAAAG	[236]	
BST	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGG	[238]	
RBS	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATGGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGTAAATT	GCTCAATGCA	TGAGGCAGCG	AAAAGAAAGG	[229]	
RMN	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATGGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCA	TGAGGCAGCG	AAAAGAAAGG	[238]	
DMT	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATGGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCA	TGAGGCAGCG	AAAAGAAAGG	[238]	
PERAM	TAACGGG - TA	CGGGGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGTAAATT	GCCCAATGCG	CGAGGCAGCG	AAGAAAAAGC	[188]	
LEIMA	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
LEIIN	TTACGGA - - A	CGGGGAATA	AGG - TC	GATT	TCGGAAGGA	GGCTGG - AAA	ATTGCTT - CA	CCTCTACCGA	AGGGCACAAG	GCCCAATG	GTCC - ATGTA	TGAGGCCGCG	AA - AGAAAAA	[229]
LEIBR	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
LEITA	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
LEPCO	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
LEPBI	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[236]	
LEPPO	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
HERMU	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGG	[237]	
HERPE	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[237]	
BLACU	TGACGGG - AG	CGGGGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGG	[238]	
TRYBR	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
TRYBE	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
TRYAV	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATC	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[237]	
TRYCR	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
TRYRA	TGACGGG - AG	CGGGGGATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATAGCTACCA	CTTCTACGGA	GGGCAGCAGG	CGCGCAAATT	GCCCAATGTA	TGAGGCAGCG	AAAAGAAAGA	[238]	
ENTO	TGACGGGTCA	CGGAGAATTA	GGGTTTGATT	CCGGAGAGGA	AGCCTGAGAC	ACGGCTACCA	TGTCCACGGT	CAGCAGCAGG	CACATAAATT	GCCCCATACG	TGAGGCAGCG	ACGGAACCGG	[238]	
PLog	TAACGAGT - A	CGGGGAATTA	GGGTTTCGACT	CCGGAGAGGG	CGCCTGAGAA	ACGGCGACCA	GATCTAAGGA	TTGCAAGAGG	CACGCAAAAT	GCCCAATGGT	TGAGGCAGTG	ACGAGAAAAAC	[238]	
PERTR	TGACGGGTAA	CAGAGAATTA	GGGTTTGATT	CTGGAGAGGG	AGCCTGAGAA	ATGGCTACCA	CTGCTATGCG	AGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	ACGAAAAAGC	[240]	
ESL	TTACGAGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGATCCAGC	[240]	
EXS	CAACGGGTGA	CGGAGAATTG	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTGCCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCA	ACGAACAAAGC	[240]	
SAC	TGACGGGTAA	CGGAGAATTT	GGGTTTCGACT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	ACGAACTAGC	[203]	
ECL	TTACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATCCG	TGAGGCAGCG	ACGAACTAGC	[240]	
TSM	TGACGGGTAA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTGCCAAGGC	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACTAGC	[240]	

	130	140	150	160	170	180	190	200	210	220	230	240	
POR	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CTGCTAAGGC	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACTAGC	[230]
TVI	TGACGGGTAA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTGCCAAGGC	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	CGAGGCAGCG	ACGAACTAGC	[229]
MAA	TGACGGGTAA	CGGAGAATTC	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	CGAGGCAGCA	ACGAACCAGC	[238]
PSSc18	TTACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACTAGC	[230]
EDS	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	CGAGGCAGCG	ACGAACTAGC	[239]
MOU	TAACGGGTAA	CGGAGAATCT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCA	ACGAACCAAC	[240]
PYR	TAACGGGTAA	CGGAGAATCT	GGGTTTCGATT	CCGGAGAGGG	AGCCTNAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCA	ACGAACCAAC	[240]
LVO	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTTCTAAGGA	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACTAGC	[239]
TEA	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCA	ACGAACTAGC	[240]
EGR	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATCCG	GGAGGCAGCG	ACGAACCAGC	[240]
EPE	CCACGGGTAA	CGGAGAATGA	GGGTTTGATA	CCGGAGAGGG	AGCCTGAGAG	ATGGCTACCA	CCGCTAAGGC	GGGCAGCAGG	CGCGCAAATT	GCCCTATACG	TGAGGCAGCA	ACGAGTCGGC	[240]
TLV	TGACGGGTGA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	ACGAACTAGC	[240]
TPL	TGACGGGTGA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	ACGAACTAGC	[240]
LBU	TGACGGGTGA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACTAAGGT	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACTAGC	[229]
DAA	CAACGGGTCA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ATGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	CGAGGCAGCA	ACGAACTAGC	[239]
LAM	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ACGGCTACCA	CTTCTAAGGA	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACCAGC	[240]
LTR	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ATGGCTACCA	CTTCTAAGGA	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACTAGC	[203]
EUGLO	TGACGGGTAA	CGAAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACAAGC	[230]
EVI	TGACGGGTGA	CGGAGAATTA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAA	ATGGCTACCA	TTACTAAGGT	GAGCAGCAGG	CGCGCAAATT	GCCCAATGCG	CGAGGCAGCG	ACGAACTAGC	[239]
LSGc38	TGACGGGTAA	CGGAGAATCA	GGGTTTGACT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTTCTAAGGA	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACTAGC	[203]
CPI	TAACGGGTAA	CGGAGAATCT	GGGTTTCGATT	CCGGAGAGGC	AGCCTGAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATCCG	TGAGGCAGCG	ACGAACTAGC	[230]
CSK	TAACGGGTAA	CGGAGAATCT	GGGTTTCGATT	CCGGAGAGGC	AGCCTGAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATCCG	TGAGGCAGCG	ACGAACTAGC	[230]
TET	TAACGGGTAA	CGGAGAATCA	GGGTTTGATA	CCGGAGAGGG	AGCCTGATAA	ACGGCTACCA	CTGCTAAGGT	GGGCAGCAGG	CACGCAAATT	GCCCAATACG	TGAGGCAGCA	ACGAGCTAGC	[230]
EGN	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	AAGAACTAGC	[240]
TVS	TGACGGGTAA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	CGAGGCAGCA	ACGAACTAGC	[240]
SNB	TGACGGGTAA	CGGAGAATTT	GGGTTTCGACT	CCGGAGAGGC	AGCCTGAGAA	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	ACGAACTAGC	[240]
EHS	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGTCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACAAGC	[240]
GRC	TGACGGGTAA	CGGAGAATCA	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CACGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACAAGC	[230]
STB	TGACGGGTAA	CGGAGAATTT	GGGTTTCGACT	CCGGAGAGGC	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	ACGAACTAGC	[240]
MUC	TGACGGGTAA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTTAGAG	ACGGCTAC - C	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACCAAC	[239]
CVL	TGACGGGTAA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACCAAC	[203]
SOA	TCACGGGTAA	CGGAGAATTT	GGGTTTCGACT	CCGGAGAGGC	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	ACGAACTAGC	[240]
SBR	TCACGGGTAA	CGGAGAATTT	GGGTTTCGACT	CCGGAGAGGC	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATACG	TGAGGCAGCG	ACGAACTAGC	[240]
CEL	TGACGGGTAA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACCAAC	[240]
CUA	TGACGGGTAA	CGGAGAATTT	GGGTTTCGATT	CCGGAGAGGG	AGCCTGAGAG	ACGGCTACCA	CTACCAAGGT	GGGCAGCAGG	CGCGCAAATT	GCCCAATGCG	TGAGGCAGCG	ACGAACCAAC	[240]

	250	260	270	280	290	300	310	320	330	340	350	360	
ACACU	AATACAATTG	GTAACAATTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGTTGTT	GCAGAAAAGC	TCGTAGTTGG	ATAAAAATTA	GAGTGAAAGC	TGGGATAATG	[359]
CAPOV	AATCCAATTG	GGAACAATTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGTTGCT	GCAGAAAAGC	TCGTAGTTGA	ACAAAAATTA	GAGTGAAAGC	TGGAATAACG	[359]
MONOV	AaTgCaAtTg	GaAcCaAtTg	GaGgGcAaGt	CtGgTgGtAa	TtCcAgCtCc	AaAgCgTaTc	TtAaGtTgTt	GcAgAaAaGc	TcGtAgTtGa	AtAaAaAtTa	GaGtGaAaGc	TgGaAtAaTg	[359]
ORNAN	AATACAATTG	GGAATCCATTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGTTGCT	GCAGAAAAGC	TCGTAGTTGG	ATAAAAATTA	GAGTGAAAGC	AGGAATGATG	[359]
NEUCR	GATACAATTG	GGAACAATTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGTTGTT	GAGGAAAAGC	TCGTAGTTGA	ACACAATCA	GATCGAAAGA	TGGAATAATA	[358]
CRYNE	AATACAATTG	GGAACAATTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGTTGTT	GCAGAAAAGC	TCGTAGTCGA	ACAAAAATTA	GAGTGAAAGC	TGGAATAATA	[358]
HARSP	AATCCAATTG	GGAACAATTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGTTGTT	GCAGAAAAGC	TCGTAGTCGA	ATAAAAATTA	GAGTGAAAGC	TGGAATAATA	[337]
ABSGL	AATGCAATTG	GGATCAATTG	AAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGTTGTT	GCAGAAACGT	CCGTAGTCGA	ATGCAATCA	GAATGAAAGT	TGGAATAATG	[358]
HEXIN	TGTCCAAGCT	ACGCAATTAG	TGGGA-GAGC	ATGGTGGTAA	TTCCATCACT	GAAGCTTTCT	CTTAGTTGTT	GCAGAAAAGC	TCGTAGCCG-	--GAAATCG	GAGTGATTAC	AAGAATGACA	[352]
TRIPY	AATGCAATTG	GTATCAATTA	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCT	AAAGCGTATA	CTAAGTTGCT	GCAGAAAAGC	TCGTAGTTGG	ATAGAAATTG	GAGTGAAAGC	TGGAATAATA	[359]
MALJA	AATGTAATTG	AT-GCAATTG	GAGGGAAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAATTTGCT	GCAGAAAAGC	TCGTAGTTGG	ATGCAATTA	GAGTGAAAGC	TGGAATAACA	[357]
MALCA	AATGTAATTG	AT-GCAATTG	GAGGGAAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAATTTGCT	GCAGAAAAGC	TCGTAGTTGG	ATGCAATTA	GAGTGAAAGC	TGGAATAACA	[357]
DINEX	GATATAATTG	CAA-CCATT-	AAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCT	AAAGTGATA	CTAAGTTGCT	GATGAAAAGC	TCGTAGTTGG	ATAGAAAATG	GAGTGAAAGC	TGGAATAATA	[357]
MONSP	AGCGGTATTG	GAACAATTA	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCT	AAAGTGATA	CTAATTTGCT	GCAGAAAAGC	TCGTAGTTGG	ATAGAAAATG	GAGTGAAAGC	TGGAATAATA	[359]
STRST	GATGCAATTG	GTACCAATTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGTGATA	CTAATTTGCT	GCAGAAAAGC	TCGTAGTTGG	ATAGAAAATG	GAGTGAAAGC	TGGAATAATA	[359]
TRIVA	CATTGG-TGG	AAACCTAGCA	GAGGGCCAGT	CTGGTGGTAA	TTCCAGCTCT	GCAGTTTGCT	CCCAATTGTT	GCAGAAACGC	CCGTAGTCT-	---AGGA---	----TA---	TATGGCTACA	[343]
SPIVO	GCTGGTGTGA	ACG-CAAGTC	GTGGGAAAAGC	ATGGTGGTAA	TTCCATCATG	ACGCAATCC	CTTAGTTGCT	GCGGCAACGC	TCGTAGCTG-	--GTAACCTCC	TAACGA----	T--GAGGGC-	[344]
ANDIN	GACGCAATCG	AAGCAACTG	GAGGGCAAAGT	CTGGTGGTAA	TACCAAGTCC	AGAGCGTATA	TTAAGTTGTT	GCAGAAAAGC	TCGTAGTTGG	ATAGAAATCA	GAGTGAAAGC	TGGGATAACC	[345]
ANDGO	GCCGCGAGTG	AAAGCAACTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AGAGCGTATA	TTAATTTGTT	GCAGAAAAGC	TCGTAGTTGG	ACAAGAATCA	GAGTGAAAGC	TGGAATAACA	[359]
SECEC	TGTTGAACGG	GTAGCAACTG	GAGGGCAAAGT	CTGGTGGTAA	CTCCAGCTCC	AGAGCGTATA	CTAATCTGTT	GCACAAAAGC	TCGTAGTCCG	ACGCAATCA	GAGTGCAAGC	TGGAATAACG	[360]
RECAM	TGTTGAACGG	GTAGCAACTG	GAGGGCAAAGT	CTGGTGGTAA	TACCAAGTCC	AGAGCGTATA	TTAAGTTGTT	GCACAAAAGC	TCGTAGTCCG	ACGCAATCA	GAGTGCAAGC	TGGAATAACG	[360]
HISAR	CTTTGAAAGG	GTAACAATTG	GAGGGCAAAGT	CTGGTGGTAA	CACCAAGTCC	AAAGCGTATA	TTAATTTGCT	GCACAAAAGC	TCGTAGTCAG	ATGCAATCA	GAGTGCAAGC	TGGAATGACG	[360]
JAKLI	TGTTGAACGG	GTATCCATTG	GAGGGCAAAGT	CTGGTGGTAA	CACCAAGTCC	AAAGCGTATA	TTAAGTTGTT	GCACAAAAGC	TCGTAGTCAG	ACGCAATCA	GAGTGCAAGC	TGGAATAACG	[360]
JAKBA	TGTACTACGG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[250]
STALI	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
NAEGR	GACTCG-GTA	GTAGCCATTG	GAGGAAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAACTGCT	GTAGAAACGC	CCGTAGTATA	CCATAAATCG	TTGTGAAAGC	TGGGACTGCA	[357]
NAEAN	GACTCG-GTA	GTAGCCATTG	GAGGAAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAACTGCT	GTAGAAACGC	CCGTAGTATA	CCGTAAATCG	TTGTGAAAGC	TGGGACTGCA	[358]
ACRRO	AGCTCG-GTA	AAA-TGACTA	GAGGAAGAGT	CTGGTGGTAA	TTCCAGCTCT	AGAGCGTATG	TTAAACTGTT	GTAGAAACGC	TCGTAGTTGG	CCACAATCA	GGGTGAAAGC	TGGGATTGTC	[357]
SAWMA	CATCCA-GTA	TAGTCTATAA	TAAAAATTGTT	GTGGT-GTAA	GTTCG-----	AAAGTCAATT	AG--TTTATG	TAGAAAAATC	ATTTTAAATAG	TTCTAAATTA	ACGCGAAAAGC	TGATGTAAACA	[347]
BST	CTCC-AGATA	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GCGCGAAAAGC	TGGGATAACA	[357]
RBS	TTTTCATTTA	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ATCAAAAGAA	GGGTGAAAGC	CGGGATAACA	[349]
RMN	TGCCCGCTTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ATCTAAAGAA	GTGCGAAAAGC	TGGGATAACA	[358]
DMT	TGCCCGCTTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ATCTAAAGAA	GTGCGAAAAGC	TGGGATAACA	[358]
PERAM	GATCCAGTTG	GTAACGATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGTTGCT	GCTGAAGGGT	TCGTAGTTGG	ACATAAAACA	GCGTGAAGGC	TGGAATGAC-	[307]
LEIMA	GTTGTA-TTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GCGTGAAAGC	TGGGATAACA	[357]
LEIIN	G-TGTA-TTT	GG-ACATTTG	TAGGAC--AT	CTTGT-GTAA	-TCCTGCTTC	AAAAGG-GTA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GCGTGAAAGC	TGGGATAACA	[341]
LEIBR	GTTGTA-TTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GCGTGAAAGC	TGGGATAACA	[357]
LEITA	GTTGTA-TTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GCGTGAAAGC	TGGGATAACA	[357]
LEPCO	GTTGTA-TTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GTGTGAAAGC	TGGGATAACA	[357]
LEPBI	GTTGTA-TTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GTGTGAAAGC	TGGGATAACA	[355]
LEPPO	GTTGTA-TTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GTGTGAAAGC	TGGGATAACA	[357]
HERMU	GTTGTA-TTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ATCTAAAGAA	GTGTGAAAGC	TGGGATAACA	[356]
HERPE	GTTGTA-TTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ATCTAAAGAA	GTGTGAAAGC	TGGGATAACA	[356]
BLACU	GTTGTA-TTT	GTAGCCATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCTAAAGAA	GTGCGAAAAGC	TGGGATAACA	[357]
TRYBR	GCCGAGTTTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCCAAAAAA	GTGCGAAAAGC	TGGGATAACG	[358]
TRYBE	GCCGAGTTTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ACCCAAAAAA	GTGCGAAAAGC	TGGGATAACG	[358]
TRYAV	GCCGAGTATT	GTAACAATTG	GAGGACAAGT	CTGATGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ATCTAAAAAA	GTGTGAAAGC	TGGGATAACA	[357]
TRYCR	GCCGAGTTTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ATCTAAAAAA	GTGTGAAAGC	TGGGATAACA	[358]
TRYRA	GCCGAGTTTT	GTAACAATTG	GAGGACAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	TTAAGCTGTT	GCTGAAGGGT	TCGTAGTTGA	ATCTAAAAAA	GTGTGAAAGC	TGGGATAACA	[358]
ENTO	GGCG-G-GAA	GGATCTACTA	GAGAGCAAAGT	CTGGTGGTAA	CTCCAGCTCT	GGAGCGTATA	CTTAGTTGCT	GCTCGAGGAC	CCTCAGTTGG	ATATGAAGCA	GAGCTTTGGC	CGGCAAGCTG	[356]
PLog	AACACGATTG	GATGCAACTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	TACAAAACGT	GTGAAGTAGC	GCTCCAAAAA	ACAAAAATTA	CCACAAAACA	GCGTGAAAAGC	TGGGATGACA	[358]
PERTR	ATTTGCTTTG	GTAACAATTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AGAGCGTATA	CTAAATTGTT	GCTGAAAAGC	TCGTAGTTCC	ATACAATCA	GTGTGAAAGC	TGAAATGACA	[360]
ESL	AATCCACTTG	GTATCAACTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	CTAATTGCT	GCTGAAACAC	CTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGAAATCTCC	[360]
EXS	AATCCACTTG	GTATCAACTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	CTAATTGCT	GCGGAAAACGT	TTGTAGTCGG	CCAAAAACCCA	GTGTGAAAGC	TGGTATGTCG	[360]
SAC	AATCCACTTG	GTAGCCACTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	CTAAATTGCT	GCGGAAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[323]
ECL	AATCCACTTG	GTAACAACTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	CTAAATTGCT	GCTGAAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCC	[360]
TSM	AATCCACTTG	GTAACGACTG	GAGGGCAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGTATA	CTAAATTGCT	GCTGAAGGAC	TTGTAGTCTG	CCAAAACTCA	GTGTGCAAGC	TGGAATGTCA	[360]

	250	260	270	280	290	300	310	320	330	340	350	360	
POR	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGTGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCGG	ACAAAACTCA	GTGTGCAAGC	TGGAATGTCA	[350]
TVI	AATTCAC TTG	GTAACGACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAGCAC	TTGTAGTCTG	CCAAAACTCA	GTGTGCAAGC	TGGAATGTCA	[349]
MAA	AATTCAC TTG	GTACCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[358]
PSSc18	AATCTACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGTGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCGG	ATAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[350]
EDS	AATTCAC TTG	GTAACAACTG	GAGGGAAAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCAGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAA -GC	TGGCATGTCA	[358]
MOU	AATCTATCTG	GTACCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGACA	[360]
PYR	AATCTATCTG	GTACCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGACA	[360]
LVO	AACTCGCTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACGC	TTGTAGTCTG	ATAAAACTCA	GTGTGAAAGC	TGGCATGGCA	[359]
TEA	AATCAACTTG	GTAACAAAGT	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	ATGCGGTATA	CTAAATTGCT	GCTGAAACAC	TCGTAGTCTG	CCTAAACTCA	GTGTGCAAGC	GTGAATTGCA	[360]
EGR	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	CTGTAGTCTG	CCAAAACTCA	GCGTGAA -GC	TGGAATGTCA	[359]
EPE	AGTTCACTCG	GCATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AATGCGGTATA	CTAAGTTGCT	GCTGAAACGC	TTGTAGTCCG	ACACAACCCA	ATGTGAAAGC	CGAGATGGCA	[360]
TLV	AATTCAC TTG	GTAACAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGGCA	[360]
TPL	AATTSACTTG	GTAACAACTG	GAGGGCAAGT	CTGGTGGTAA	TT -CAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGGCA	[358]
LBU	AATTCAC TTG	GTATCGACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	ATTGCGGTATA	CTAAATTGCT	GTGAAACAC	TTGTAGTCTG	ATAA -ACTCA	GTGTGAAAGC	TGGAATGTCA	[348]
DAA	AATGCAC TTG	GAATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGGTATG	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGACG	[359]
LAM	AATTCAC TTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AACGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	ATAAAACTCA	GTGTGAAAGC	TGGAATGTCTG	[360]
LTR	AATTCAC TTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGGTATA	CTGAGTTGCT	GCTGAAACAC	TTGTAGTCTG	ATAGAACTCA	GTGTGAAAGC	TGGAATGTCA	[323]
EUGLO	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGNCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAACGTCA	[350]
EVI	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	CTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATTTCA	[359]
LSGc38	AAGCCCTTG	GTATCAACCG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	GATGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	ATAAAACTCA	GTGTGCAAGC	TGGAATGTCA	[323]
CPI	AATCCACTTG	GTACCTATTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CTAAAACTCA	GTGTGAAAGC	TGGAACGTCA	[350]
CSK	AATTCAC TTG	GTACCTATTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAACGTCA	[350]
TET	AATCCATTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACGC	TTGTAGTCGA	CCACAACCTCA	GAGTGAAAGC	TGGGATGACA	[350]
EGN	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTATATTGCT	GCTGAAACAC	CTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAACGTCA	[360]
TVS	AATTCAC TTG	GTAACGACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAACGC	TGGAATGGCA	[360]
SNB	AATTCAC TTG	GTAGCCACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCGGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[360]
EHS	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CTAAAACTCA	GTTTGAAGA	TGGAATGTCA	[360]
GRC	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAAGCGGTATA	CTAAATTGCT	GCTGAAACAC	TTGTAGTCTG	CTAAAACTCA	GTTTGAAGA	TGGAATGTCA	[350]
STB	AATTCAC TTG	GTAGCCACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCGGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[360]
MUC	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATG	CTAAATTGCT	GCGGAAACAC	TTGTAGTCTG	CCATAACTCA	GTGTGAAAGC	TGGAATGTCA	[359]
CVL	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATG	CTAAATTGCT	GCGGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[323]
SOA	AATTCAC TTG	GTAGCCACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCGGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[360]
SBR	AATTCAC TTG	GTAGCCACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATA	CTAAATTGCT	GCGGAAACAC	TTGTAGTCCG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[360]
CEL	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATG	CTAAATTGCT	GCGGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[360]
CUA	AATCCACTTG	GTATCAACTG	GAGGGCAAGT	CTGGTGGTAA	TTCCAGCTCC	AAGGCGGTATG	CTAAATTGCT	GCGGAAACAC	TTGTAGTCTG	CCAAAACTCA	GTGTGAAAGC	TGGAATGTCA	[360]

	370	380	390	400	410	420	430	440	450	460	470	480	
ACACU	AAATAGA---	ATGATGGCAT	TAATATTTAA	TTGTCAGAGG	TGAAATTCTT	GGATTTATGA	AAGATTAAC	TCTGCGAAAG	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTTAGGGGA	[476]
CAPOV	AAATAGA---	ATGATGCCGT	TAGTATTTAA	TTGTCAGAGG	TGAAATTCTT	GGATTTATGA	AAGACTAACT	TCTGCGAAAG	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTTAGGGGA	[476]
MONOV	GaAtAgA---	AtGaTgGcAt	TaGtAtTtAa	TtGtCaGaGt	TgAaAtTcTt	GgAtTtAtGa	AaGaCtAaCt	TcTgCgAaAg	CcAaGgAtGt	TtTcAtTaAt	CaAgAaCgAa	AgTtAgGgGa	[476]
ORNAN	GAATAGA---	ATGATGGCAT	TCGTATTGTG	CGACCGGCGC	TGAAATTCTT	GGACCGGCGC	AAGACGGACC	AAGCGAAAG	CCAAGATGT	TTTCATTAAT	CAAGAACGAA	AGTCGGAGGT	[476]
NEUCR	GAATAGA---	ATGATGGCAT	CAGTATTCAA	TTGTCAGAGG	TGAAATTCTT	GGATTTATTG	AAGACTAACT	ACTGCGAAAG	CCAAGGATGT	TTTCATTAAT	CAGGAACGAA	AGTTAGGGGA	[475]
CRYNE	GAATAGA---	ATGATGGCAT	TGGTATTCCG	TTGTCAGAGG	TGAAATTCTT	AGATTGACGG	AAGACCAACA	ACTGCGAAAG	CCAAGGACGT	TTTCATTGAT	CAAGAACGAA	GGTTAGGGGA	[475]
HARSP	GAATAGA---	ATGATGGCAT	TAGTATTTAA	TTGTCAGAGG	TGAAATTCTT	GGATTTATGA	AAGACTAACT	TCTGCGAAAG	CCAAGGATGT	TTTCATTGAT	CAAGAACGAA	AGTTAGGGGA	[454]
ABSG	AAATAGA---	ATGATGGCAT	TTGTATTGG	CGCTAGAGG	TGAAATTCTT	GGATTGGCCG	AAGACAAACT	ACTGCGAAAG	CCAAGGACGT	TTTCATTGAT	CAGGGACTAA	AGTTGAGGGA	[475]
HEXIN	TTGC--T---	GGAGTCTGA-	TAGTATTGG	TGGGGACAGG	TGAAATAGGA	TGATCCACCA	AAGACTTTCA	ACAGCGAAGG	CAAGCGCCGT	CC-AGYTG	CAAGAGCGAA	AGTGTGAGGA	[465]
TRIPY	AAATAGA---	ATGATGGCAT	TTGTATTGG	TCGTCAGAGG	TGAAATTCTT	GGATTGACGC	AAGACAAACT	ACTGCGAAAG	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTTAGGGGA	[476]
MALJA	AAATA-G---	ATGGTGGCAT	TCGTACCTTG	GAGTCAGAGG	TGAAATTCTT	AGATTTC	GAGACGAACT	TCTGCGAAAG	CCAAGGATGT	TCTCAYTAAT	CAAGAACGAA	AGTTAGGGGA	[473]
MALCA	AAATG-----	ATGGTGGCAT	TAGTACCTTG	GAGTCAGAGG	TGAAATTCTT	AGATTTC	GAGACTAACT	ACTGCGAAAG	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTTAGGGGA	[472]
DINEX	AAATAGA---	ATGATGGCAT	TAGTATTGCA	TTGTCAGAGG	TGAAATTCTT	GGATTTATGC	AAGACTAACT	AGTGC	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTGAGGGGA	[474]
MONSP	AAATAGACAC	ATGATGGTAT	CTGTATTGCA	TAGTCAGAGG	TGAAATTCTT	GGATTTATGC	AAGACAAACA	AGTGC	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTTAGGGGA	[479]
STRST	AAATAGA---	ATGATGGCAT	GGGTATTGCA	TGAAATTCTT	AGATTTC	AAGACTTTCA	AAGACTTTCA	ACTGCGAAAG	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTTAGGGGA	[476]
TRIVA	TGA-----	AAGCTGGTAG	ATCTATCTCA	TGACGAGTGG	TGGAATACTT	TGACTCATGA	GAGAGAAGCT	GAGGCGAAGG	CCTAGAGGGT	TTCTGTCGAT	CAAGGGCGAG	AGTAGGAGTA	[456]
SPIVO	-----TA---	-TGCCCA-	CTCTACTAGC	GGAGGACAGG	TGAAATAGCA	TGATTCTCT	AAGAGGATCG	GCTGCGTAAG	CGACT-GTGC	TCTTGCTGGT	CATGAGCGAG	AGCGCAGGGC	[451]
ANDIN	ACCCTGA---	TTGATGGCAT	TCGTATTCCG	AAGTCAGAGG	TGAAATTCTT	GGATTTC	AAGACTTTCA	TATGCGAAGG	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTTGCGGGA	[462]
ANDGO	ACGAAGA---	CTGATGGCAT	TCGTATTCCG	TAGTCAGAGG	TGAAATTCTT	GGATTTC	AAGACGAACT	TATGCGAAGG	CCAAGGATGT	TTTCATTAAT	CAAGAACGAA	AGTTGCGGGA	[476]
SECEC	AAGCGGA---	ATGGTGGCAT	TCGTATTCCG	CGCTCAGAGG	TGAAATTCTA	GGATTGCGG	AAGACGAACT	ACTGCGAAGG	CCAAGGATGT	TCCCATTTGAT	CAAGAACGAA	GTCTAGGGGA	[477]
RECAM	ACGCGGA---	ATGGTGGCAT	TTGTATTCC	AGTCAGAGG	TGAAATTCTA	GGATTTC	AAGACGAACT	ACTGCGAAGG	CCAAGGATGT	TCCCGTTGAT	CAAGAACGAA	AGTTAGGGGA	[477]
HISAR	AAACGGA---	ATGGTGGCAT	TCGTATTCCA	ACGTCAGAGG	TGAAATTCTA	GGATTGTTG	AAGACGAACT	GCTGCGAAGG	CCAAGGATGT	TCCCGTTGAT	CAAGAACGAA	AGTTAGGGGA	[477]
JAKLI	ATGCGGA---	ACGCTGGCAT	TCGTATTCCG	GTGTCAGAGG	TGAAATTCTA	GGATTACCTG	AAGACGAACT	ACTGCGAAGG	CCAAGGATGT	TCCCATTTGAT	CAAGAACGAA	AGTTAGGGGA	[477]
JAKBA	-----	-TGCCAT	TCGTATTCCG	TGCTAGAGG	TGAAATTCTA	GGATTACCTG	AAGACGAACT	AATGCGAAGG	CCAAGGATGT	TCCCATTTGAT	CAAGAACGAA	AGTTAGGGGA	[356]
STALI	-GCTAGC---	-GTGGGCGAC	AAGTATTCCG	GTGACAGGGG	TGAAATCCGT	ACTCCCGGAA	GAC--TACCA	GATGCGAAGG	CCCACGGCGT	TTCCGTTGAT	CAAGAACGAA	AGTTAGGGGA	[113]
NAEGR	GAGTA-----	GCTTGGCCCC	GGGTACCATG	AGGCTAGAGG	TGAAATTCTG	AGACCCCTCAT	GTGACCAACT	AAGGCGAAGG	CCCACACAA	GCTCGCTCAT	CAGGGACAAA	AGTTGGGGGA	[472]
NAEAN	GAGAA-----	GCTTGGCCCC	GGGTACCATG	AGGCTAGAGG	TGAAATTCTG	AAACCCCTCAT	GTGACCAACT	AAGGCGAAGG	CCCACACAA	GCTCGCTCAT	CAGGGACAAA	AGTTGGGGGA	[473]
ACRRO	CTATA-----	GCTCGGGCTC	TGGAACATA	GTGCTAGAGG	TGAAATTTAG	ATACCGCTAT	ATGCTATCT	GATGCGAAGG	CCCAGTACAT	TTCAATCGAT	CAGGAACGAA	AGTTAGGGGA	[472]
SAWMA	TACATAG---	-TAGAGTTCT	CAGTACTTGC	TAACTAGGGG	TGAAATCCGT	ATAGTTAGCA	AAGACTACCA	AAGGCGAAGG	CCCAGAAGTG	GTCTATTGAT	CAGGAGCTAG	AGTTAAGGGA	[463]
BST	AAGGACAAC	GGAATGGAGA	ACGTACTGGA	CGCTGAGAGG	TGAAATTCTT	AGACCGTTCC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[477]
RBS	AAGGA--CC	GGAATGGAGA	ACGTACTTGA	CGCTGAGAGG	TGAAATTCTT	AGACCGCTCA	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[466]
RMN	AAGCACAGCC	GGAATGGAAA	ACGTACTTGG	CGCTGAGAGG	TGAAATTCTT	AGACCGGCCA	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[478]
DMT	AAGGACAGCC	GGAATGGAAA	ACGTACTTGA	CGCTGAGAGG	TGAAATTCTT	AGACCACTCA	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[478]
PERAM	--AGA-GCA	-GGATGAAAA	CCGTATTCCG	CGCTGAGAGG	TGAAATTCTT	AGACCGCCCG	GGGACGAACT	CAAGCGAAGG	CCAAGTATAC	CTTCATCAAT	CAAGAACCAA	AGTGTGGAGA	[422]
LEIMA	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[477]
LEIIN	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[461]
LEIBR	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[477]
LEITA	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[477]
LEPCO	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[477]
LEPBI	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[475]
LEPPO	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[477]
HERMU	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTGAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[476]
HERPE	AAGGACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTGAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CAAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[476]
BLACU	AGAGACAGCC	GGAATGGAGA	ACGTACTCGA	CGCTCAGAGG	TGAAATTCTT	AGACCGATCG	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[477]
TRYBR	AAGCACAGCC	GGAATGGAGA	ACGTACTGGG	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[478]
TRYBE	AAGCACAGCC	GGAATGGAGA	ACGTACTGGT	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[478]
TRYAV	AAGGACAGCC	GGAATGGAGA	ACGTACTGGT	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[477]
TRYCR	AAGGACAGCC	GGAATGGAGA	ACGTACTGGT	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[478]
TRYRA	AAGGACAGCC	GGAATGGAGA	ACGTACTGGT	CGCTCAGAGG	TGAAATTCTT	AGACCGCACC	AAGACGAACT	ACAGCGAAGG	CCAAGGATAC	CTTCCTCAAT	CAAGAACCAA	AGTGTGGAGA	[478]
ENTO	ATTCGAC---	GGTATTTTGT	CTGTACCTTC	AGGGAAGCGG	TGATATGCGA	TGATCTCTGAA	GGGACAACCT	AACGCGAAGG	CACTGAACGG	CTTAATCTAT	CATGAGTCTG	AGCATGAGGA	[473]
PLog	TTGGGATGAC	-GGATGAGAT	ACGTATTGTG	CGGCTAGTGG	TGAAATAGCA	AGACCGTCA	AGACCGAACA	TCTGCGAAGG	CCAAGGATGC	CTTCCTCAAT	CAAGAGCGAT	AGTGTGGGA	[477]
PERTR	AGGGACGAC-	GTTATGTGGT	GTGTATTTCT	CGGTTAGAGG	TGAAATTCTT	AGACCGTAGA	AGGACATACT	GCGGCGAAGG	CCAAGTCTAG	CTCTGTCTAT	CAAGAACGAA	AGTGTGGGA	[479]
ESL	AAGCACGAC-	GAATCGTGAA	AGATACTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGATCAGCG	GNGGCGAAGG	CCAAGTGCTC	GTTCTGCTGAT	CAAGGATGAG	AGTCTGGGA	[479]
EXS	TCGCACGACT	GCCGCGTGCA	AGGAATGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	GCGGCGGACA	GCGGCGGACA	CCAAGGCGCG	GTTCTGCTGAT	CAAGAGCGAG	AGTCTGGGA	[480]
SAC	AAGCACGACC	GCCAGGTGGA	AGGTATTGGA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACT	GCAAGCGAAGG	CCAAGAGCAT	GTTCTGCTGAT	CAAGAATGAG	AGTCTGGGA	[443]
ECL	CAGCACGACC	GGAATGTGGA	CGATATTGTA	GCGCCAGAGG	TGAAATTTTC	AGATCGCTGC	AAGATCCACG	GCAAGCGAAGG	CCTAGTGCGT	GTTCTGCTGAT	CAAGAATGAG	AGTATGGGA	[480]
TSM	AAGCACGACC	GAAAGAGCGGA	AGGTACTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACT	GCAAGCGAAGG	CCAAGTGATC	GTTCTGCTGAT	CAAGAATGAR	AGTGTGGGA	[480]

	370	380	390	400	410	420	430	440	450	460	470	480	
POR	AAGCACGACC	GGCCAGTGAA	GGATATGGCT	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	CAGATCAACT	GCAGCGAAGG	CCAAGTGCAT	GTTTCGTCGAT	CAAGAATGAG	AGTGTGGGGA	[470]
TVI	AAGCACGACC	GAAGAGCGAA	AGGTATGGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	TGGACCAACT	GCAGCGAAGG	CCAAGTGCAC	GTTTCGTCGAT	CAAGAATGAG	AGTGTGGGGA	[469]
MAA	AAGCACGACC	GGAGAGTGCA	AGATATTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	ACGATCCACA	GCAGCGAAGG	CCAAGGATGA	GTTTCGTCGAT	CAAGAATGAG	AGCCTGGGGA	[478]
PSSc18	AAGTACGACC	GACCAAGTGAA	GGGTATTGTA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACT	GCAGCGAAGG	CCAAGTGCAT	GTTTCGTCGAT	CAAGAATGAG	AGTGTGGGGA	[470]
EDS	AGGCACGACC	GAATCGTGAA	AGGTACTACC	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCCACG	GCGGCGAAGG	CCAAGGCTGT	GTTTCGTCGAT	CAAGAATGAG	AGTCCGGGGA	[478]
MOU	AAGCACGACC	GGAGAGTGAA	AGATATTGCA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGATCCCTT	GCGGCGAAGG	CCATGTCTGG	ATTTCGTCGAT	CAAGAATGAG	AGCCTGGGGA	[480]
PYR	AAGCACGACC	GGAGAGTGGA	AGGTATTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCCACT	GCGGCGAA - G	CCATGTTTGA	GTTTCGTCGAT	CAAGAATGAG	AGCCTGGGGA	[479]
LVO	AAGCACGATC	GCCTGGTGAA	AGATATTGCA	ACGCCAGAGG	TGAAATTCTT	AGATCGCTGC	AAGATGACCC	GCAGCGAAGG	CCAAGTGCTC	GTTTCGTCGAT	CAAGAACGTG	AGTGTGGGGA	[479]
TEA	GAGCACGATC	GGACAGTGGA	AGATACTGCA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGATCAACT	GCAGCGAAGG	CCAAGAGCTC	GTTTCGTCGAT	CAAGAATGAG	AGTGTGGGGA	[480]
EGR	GAGCACGACC	GTCCTGTGAA	AGATACTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGATCGACC	GCAGCGAAGG	CCAAGCGCGT	GTTTCGTCGAT	CAAGGATGAG	AGTCCGGGGA	[479]
EPE	ATGGACGAC -	GGGATGTGCT	GGATATTGAC	ACGCCAGAGG	TGAAATTCTT	AGACCGCAGT	GGGATCGACT	GCGGCAAAGG	CCACGATTCT	GTTTCGTCGAT	CAAGAACGAG	AGTGTGGGGA	[479]
TLV	AAGCACGACC	GCATCGTGAA	AGGTATTGCA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	CAGACCAACT	GCAGCGAAGG	CCAAGTGCAT	GTTTCGTCGAT	CAAGAATGAG	AGCCTGGGGA	[480]
TPL	AAGCACGACC	GCATCGTGAA	AGGTATTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	CAGACCAACT	GCAGCGAAGG	CCAAGTGCAT	GTTTCGTCGAT	CAAGAATGAR	ASCCTGGGGA	[478]
LBU	AAGCACGACT	GTCCAGTGAA	AGGTATTGCA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACT	GCAGCGAAGG	CCAAGTGCAA	GTTTCGTCGAT	CAAGAATGTG	AGTGTGGGGA	[468]
DAA	AAGCACGACC	GGAGAGTGAA	CGATATTGAA	GCGCTAGAGG	TGAAATTCTC	AGACCGCTGC	AAGATCCACT	GCAGCGAAGG	CCAAGTGCTT	GTTTCGTCGAT	CAAGAATGAG	AGTGTGGGGA	[479]
LAM	CAGTACGACC	GTCCTGTGAA	AGATATTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGATCAACT	GCGGCAAAGG	CCAAGTGCAT	GTTTCGTCGAT	CAAGAATGTG	AGTGTGGGGA	[480]
LTR	TAGCATGACC	ACTCTGTGGA	AGGTATTGCA	ACGCCAGAGG	TGAAATTCTC	AGATCGCTGC	AAGACCAACT	GCGGCGAAGG	CCAAGTGCAT	GTTTCGTCGAT	CAAGAATGTG	AGTGTGGGGA	[443]
EUGLO	AGGCACGACC	GTAACGTGAA	AGATATGGGA	GCGCCAGAGG	TGAAATTCTT	AGATCGCTGC	CAGATCCACT	GCAGCGAAGG	CCAAGTGCAC	GTTTCGTCGAT	CAAGAATGAG	AGTTCGGGGA	[470]
EVI	AAGCACGACC	GGAGAGTGAA	AGGTATTGCA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACT	GCAGCGAAGG	CCAAGTGCAT	GTTTCGTCGAT	CAAGGATGAG	AGTCCGGGGA	[479]
LSGc38	AAGCACGACC	GTCCAGTGGA	AGGTATTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACG	GCGGCGGAGG	CCAGGTGCAC	GTTTCGTCGAT	CAAGAATGTG	AGTGTGGGGA	[443]
CPI	CAGCATGACC	AGAGAGTCAA	AGATATTGCA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AGGATCAACA	GCAGCGAAGG	CCAAGTGCAT	GTTTCATCGAT	CAAGAATGAA	AGCGTGGGGA	[470]
CSK	CAGCATGACC	AGAGAGTCAA	AGATATTGCA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AGGATCAACA	GCAGCGAAGG	CCAAGTGCAT	GTTTCATCGAT	CAAGAATGAA	AGCGTGGGGA	[470]
TET	AAGGACGACC	GGCAGGTGAT	GGGTATTGAC	GCGCTAGAGG	TGAAATTCTT	AGATCGCTGC	GGGACCTTCT	GCAGCGAAGG	CCAAGGGCAT	GTTTCGTCGAT	CAAGAACGAG	AGTGTGGGGA	[470]
EGN	AAGCACGACC	GAATCGTGAA	AGGAAGTGCA	GCGCCAGAGG	TGAAATTCTT	AGATCGCTGC	AAGACCCACT	GCAGCGAAGG	CCAAGGACGT	GTTTCGTCGAT	CAAGGATGAG	AGTCTGGGGA	[480]
TVS	AAGCACGACC	GAATCGTGAA	AGGTATTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACT	GCAGCGAAGG	CCAAGTGCAT	GTTTCGTCGAT	CAAGAATGAG	AGCGTGGGGA	[480]
SNB	AAGCACGACC	GCCAAGTGGA	AGGTATTGGA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACT	GCAGCGAAGG	CCAAGAGCAT	GTTTCGTCGAT	CAAGAATGAG	AGCTCGGGGA	[480]
EHS	AGGCACGACC	GTACAGTGAA	AGATATGGGA	GCGCCAGAGG	TGAAATTCTT	AGATCGCTGC	CAGATCCACT	GCAGCGAAGG	CCAAGTGCAC	GTTTCGTCGAT	CAAGAATGAG	AGTTCGGGGA	[480]
GRC	AGGCACGACC	GTACAGTGAA	AGATATGGGA	GCGCCAGAGG	TGAAATTCTT	AGATCGCTGC	CAGATCCACT	GCAGCGAAGG	CCAAGTGCAC	GTTTCGTCGAT	CAAGAATGAG	AGTTCGGGGA	[470]
STB	AAGCACGACC	GAAGAGTGGA	AGGTATTGGA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGC	AAGACCAACT	GCAGCGAAGG	CCAAGGGCAT	GTTTCGTCGAT	CAAGAATGAG	AGCGTGGGGA	[480]
MUC	AAGCACGACC	GGAGAGTGAA	AGATATTGCA	ACGCTAGAGG	TGAAATTCTT	AGACCGTTGC	AAGATCAACG	GCAGCGAAGG	CCAAGTGCCT	GTTTCGTCGAT	CAAGAATGAG	AGCGTGGGGA	[479]
CVL	AAGCACGACC	GGAGAGTGAA	AGATATTGCA	ACGCTAGAGG	TGAAATTCTT	AGACCGCTGC	AAGATGAACG	GCAGCGAAGG	CCAAGTGCCT	GTTTCGTCGAT	CAAGAATGAG	AGCGTGGGGA	[443]
SOA	AAGCACGACC	GAAGAGTGGA	AGGTATGTGA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGA	CAGACCAACT	GCAGCGAAGG	CCAAGGGCAT	GTTTCGTCGAT	CAAGAATGAG	AGCGTGGGGA	[480]
SBR	AAGCACGACC	GAAGAGTGGA	AGGTATGTGA	GCGCCAGAGG	TGAAATTCTT	AGACCGCTGA	CAGACCAACT	GCAGCGAAGG	CCAAGGGCAT	GTTTCGTCGAT	CAAGAATGAG	AGCGTGGGGA	[480]
CEL	AAGCACGACC	GGAGAGTGAA	AGATATTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGTTGC	AAGATCAACG	GCAGCGAAGG	CCAAGTGCCT	GTTTCGTCGAT	CAAGAATGAG	AGCGTGGGGA	[480]
CUA	AAGCACGACC	GGAGAGTGAA	AGGTATTGCA	ACGCCAGAGG	TGAAATTCTT	AGACCGTTGC	AAGACCAACG	GCAGCGAAGG	CCAAGTGCCT	GTTTCGTCGAT	CAAGAATGAG	AGCGTGGGGA	[480]

	490	500	510	520	530	540	550	560	570	580	590	600	
ACACU	TCGAAGACGA	TCAGATACCG	TCGTAGTCTT	CCATAAACGA	TGCCGACTGA	GAAATCGTCT	TTGGGTTCCG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[596]
CAPOW	TCGAAGACGA	TCAGATACCG	TCGTAGTCTT	CCTTAAACGA	TGCCCAACTGA	GAAATCGTTT	TTGGGTTCTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGC-	[595]
MONOV	TcGaAgAcGg	TtAgAtAcCg	TcGtAgTcCt	CcAtAaAcTa	TgCcGaAcTg	GaAaTcGtCt	TtGgGtTcCg	GgGgGtGtAt	GgTcGcAaGg	CtGaAaCtTa	AaGgAaTtGa	CgGaAgGgCa	[596]
ORNAN	TCGAAGACGA	TCAGATACCG	TCGTAGTTCC	CCATAAACGA	TGCCGACCGG	GAAACCGTCT	TTGGGTTCTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[596]
NEUCR	TCGAAGACGA	TCAGATACCG	TCGTAGTCTT	CCATAAACTA	TGCCGATCGA	TAAATCATGT	TTGGGCTCCT	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGAAATTGA	CGGAAGGGCA	[595]
CRYNE	TCAAAACGA	TTAGATACCG	TTGTAGTCTT	CAGTAAACGA	TGCCGACCGA	GAAATCGTCT	TTGGGTTCTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[595]
HARSP	TCGAAGACGA	TCAGATACCG	TCGTAGTCTT	CCATAAACTA	TGCCAGTGA	GAAATCGCTT	??????????	??????????	??????????	??????????	??????????	??????????	[574]
ABSGI	TCGAAGACGA	TTAGATACCG	TCGTAGTCTT	CCACAACTA	TGCCGAC-GC	GAAAGTGTCT	TTGGGTTCTG	GGGGGAGTAT	GGACGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[594]
HEXIN	TAGACGATGA	TTAGAAACCG	TTTTATTCCA	CCTTAAACGA	TACCATCAGG	GAAACCGGTT	TCGGGTTCTG	GGGGAAGTAT	GATTGCAAGA	TTGAAACTTG	AAGGTATTGA	CGGAAGGATA	[585]
TRIPY	TCGAAGGCGA	TTAGAAACCG	CCCTAGTCTT	CCATAAACGA	TGCCAACAGA	GAAATCGTCT	CTGGGTTCTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGAAATTGA	CGGAAGGGCA	[596]
MALJA	TCGAAGACGA	TCAGATACCG	TCGTAGTCTT	CCATAAACGA	TACCGACTGA	GAAATCGTCT	TTGGGTTCTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[593]
MALCA	TCGAAGACGA	TCAGATACCG	TCGTAGTCTT	CCATAAACGA	TACCGCCTGA	GAAATCGTCT	TTGGGTTCTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[588]
DINEX	TCGAAGACGA	TCAGAAACCG	TCGTAGTCTT	CCATAAACTA	TACCGTCAAG	GAAACCGTTT	TCGGGTTCTG	GGGGGAGTAT	GGCCGCAAGA	CTGAAACTTA	GAGAAATTGA	CGGAAGGGCA	[594]
MONSP	TCGAAGGCGA	TCAGAAACCG	CTCTAGTCTT	CTATAAACGA	TGCCAACAGA	GAAATCGTCT	CTGGGTTCTG	GGGGGAGTAT	GGTCGCAAGT	CTGAAACTTA	AAGAAATTGA	CGGAAGGGCA	[599]
STRST	TCGAAGGCGA	TCAGAAACCG	CTCTAGTCTT	CTATAAACGA	TGCCAACAGA	GAAATCGTCT	CTGGGTTCTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[596]
TRIVA	TCCAACAGGA	TTAGAGACCC	TGGTAGTTCC	CCTTAAACGA	TGCCGACGGA	GAAATCGTTT	CTGGGTTCTG	GGGGAAGTAT	GACCGCAAGG	CTGAAACTTG	AAGGAATTGA	CGGAAGGGCA	[576]
SPIVO	TGGACGACGA	TCAGAGACCG	TTCTACTCTG	CCCTAAACGA	TACTGCC-GA	AGTCCTGTCT	TCGAGTTCTT	GGGGGAGTAT	GGTCGCAAGA	GTGAAACTTG	AAGGTATTGA	CGGAGGC-TA	[569]
ANDIN	TCGAAGACGA	TCAGATACCG	TCGTAGTCCG	CCATAAACGA	TGCCGACGGA	GAAACCGTCT	TTGGGTTCCG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[582]
ANDGO	TCGAAGATGA	TTAGATACCG	TCGTAGTCCG	CCATAAACGA	TGCCGACTGA	GAAATCGTCT	TTGGGTTCCG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[596]
SECEC	TCGAAGACGA	TTAGACACCG	TCATAGTCTT	AGGTAACGA	TGCCGACTGG	GAAACCGTCT	CTGGGTTGTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[597]
RECAM	TCGAAGACGA	TCAGACACCG	TCCTAGTCTT	CCGTAACGA	TGCCGACCGG	GAAACCGTCT	TTGGGTTGTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[597]
HISAR	TCGAAGACGA	TTAGACACCG	TCCTAGTCTT	CCGTAACGA	TGCCGACCGG	GAAACTGTCT	TTGGGTTGTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[597]
JAKLI	TCGAAGACGA	TCAGACACCG	CCCTAGTCTT	CCATAAACGA	TGCCGACAGG	GAAACTGTCT	TTGGGTTGTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[597]
JAKBA	TCGAAGACGA	TCAGACACCG	TCCTAGTCTT	CCATAAACGA	TGCCGACAGG	GAAACTGTCT	TCGGGTTGTG	GGGGGAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[476]
STALI	TCGAAGATGA	TTAGATATCG	TCGTAGTCTT	CCGTAACGA	TACTGCCAGG	GAAACCGTTT	TCGAGTTCTG	GGGGAAGTAT	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[233]
NAEGR	TCGAAGACGA	TTAGATACCG	TCGTAGTCCC	CTATAAACGA	TACCAACAGG	GAAACCGTCT	TTGGGTTCTG	GGGGGAGTAT	AGTCGCAAGA	CCGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[592]
NAEAN	TCGAAGACGA	TTAGATACCG	TCGTAGTCCC	CTATAAACGA	TACCGACAGG	GAAACCGTCT	TTGGGTTCTG	GGGGGAGTAT	AGTCGCAAGA	CCGAAACTTA	AAGGAATTGA	CGGAAGGGCA	[593]
ACRRO	TCGAAGATGA	TTAGATACCG	TCGTAGTCTT	CTATAAACGA	TACCAACAGG	GAAACCGTTT	TTGGGTTCTG	GGGGAGTTAT	AGTCGCAAGG	CCGAAACTTA	AAGSAATTGA	CGGAAGGGCA	[592]
SAWMA	ACAAAGATGA	TTAGATACCG	TCGTAGTCTT	CTATAAACTA	TACTAATCGA	GAAATCAATT	TTGAGTCTCA	GGGGAAGTAT	ATTGCGAAGA	ATGAAACTTA	AAGGAATTGA	CGGAATTCTA	[583]
BST	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CCACAACCGG	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTGCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
RBS	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CCACAACCGG	GGACCCCGCG	AAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[586]
RMN	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CCACAACCGG	TGACCCCGCG	GAAAGCGGTT	ACAGTCTCAG	GGG-GAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
DMT	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CCACAACCGG	TGACCCCGCG	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[598]
PERAM	TCGAAGATGA	TTAGATACCA	TTGTAGTCCA	CTGCAACCGA	TGACGCCCGC	GAAAGCGGCT	ATGGTCTCAT	GGGGAAGTAC	GTTCGCAAGG	CTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[542]
LEIMA	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
LEIIN	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[581]
LEIBR	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
LEITA	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
LEPCO	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
LEPBI	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[595]
LEPPO	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
HERMU	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCG	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[596]
HERPE	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCG	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[596]
BLACU	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
TRYBR	TCAAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[598]
TRYBE	TCAAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[598]
TRYAV	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[597]
TRYCR	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[598]
TRYRA	TCGAAGATGA	TTAGAGACCA	TTGTAGTCCA	CTGCAAACCGA	TGACACCCGC	GAAAGCGGTT	ACAGTCTCAG	GGGGGAGTAC	GTTCGCAAGA	GTGAAACTTA	AAGAAATTGA	CGGAATGGCA	[598]
ENTO	TCGAAAAGGA	TCAGATACCG	TTGTAGTCCG	CCCTAAACGA	TAGTGCCAGT	GAAACCGGTT	CTGCCCTCAG	GGGGAAGTAT	AGTCGCAAGT	CTGAAACTTA	AAGGAATTGA	CGGACTCGCA	[593]
PLoG	TCTAAGATGA	TCAGATACCG	TTGTAGTCCA	CTGTAACCGC	TGCCAGCTGA	GAAATCGCC-	-TGGGTTCCG	GGGGGAGTAC	GTTCGCAAGG	CCGAAACTTA	AAGGAATTGA	CGGAATGGCA	[595]
PERTR	GCAAAGATGA	TCAGATACCG	TTGTAGTCCA	CTGTAACCGA	TGTCAG-GCGA	GAAATCACCT	ATGGATTCTG	GGGGGAGTAC	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAGTGACA	[598]
ESL	GCAAAGAGGA	TCAGACACCG	TCGTAGTCCA	CTGTAACCGC	TGCCAGCCGA	GAAATCGCCT	GTGGGTTCTG	GGGGGAGTAC	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[599]
EXS	TCGAAGAGGA	TCAGACACCG	TCGTAGTCCA	CTGTAACCGA	TGCCGCGCGA	GAAATCGCCT	TTGGGTTCCG	GGGGGAGTAC	GTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
SAC	GCGAAGATGA	TCAGACACCG	TCGTAGTCCG	CTGTAACCGA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTCCG	GGGGGAGTAC	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[563]
ECL	GCAAAGATGA	TCAGACACCG	TCGTAGTCCA	CCGTAACCGA	TGCCGCGCGA	GAAATCGCCT	GTGGGTTCTG	GGGGGAGTAC	GGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
TSM	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAACCGA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTCTG	GGGGGAGTAC	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]

	490	500	510	520	530	540	550	560	570	580	590	600	
POR	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACGA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[590]
TVI	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACGA	TGCCAGCCGA	GAAATCGCCA	TTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[589]
MAA	TCCAAGATGA	TTAGACACCG	TCGTAGTCCA	CTGTAAACCA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCA	G	CGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[598]
PSSc18	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[590]
EDS	GCAAAGATGA	TCAGACACCG	TCGTAGTCCG	CTGTCAACCG	TGCCAGCCGA	GAAATCGCCC	ATGGGTTCTG	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[598]
MOU	GCAAAGGTGA	TTAGACACCT	CCGTAGTCCA	CCGTAAACCA	TGCCGGCCGA	GAAATCGCCT	TTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
PYR	GCAAAGGTGA	TTAGACACCT	CCGTAGTCCA	CTGTAAACCA	TGCCGGCCGA	GAAATCGCCT	TTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[599]
LVO	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACGA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[599]
TEA	ACAAAGATGA	TCAGACACCG	TCGTAGTCCA	CAGTAAACCA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCA	G	TGTCGCAAGG	CAGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
EGR	GCAAAGATGA	TCAGACACCG	TCGTAGTCCG	CCGTAAACCG	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCA	G	GTTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[599]
EPE	GCAAAGATGA	TCAGATACCG	TTGTAGTCCA	CCGTAAACCG	TGTGAGCCGA	GAAATCGCCT	GTGGATTTCAG	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGATTGGCA	[599]
TLV	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACCA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
TPL	GCARAAATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACCA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CCGAATGGCA	[598]
LBU	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACCA	TGCCAGCCGA	GAAATCGCCC	ATGGGTTCAA	G	TGTTGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[588]
DAA	GCAAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTCAACTG	TGCCGGCCGA	GAAATCGCCA	TTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[599]
LAM	GCAAAGATGA	TCAGACACCG	TTGTAGTCCA	CTGTAAACCA	TGCCAGCCGG	GAAACCGCCT	ATGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
LTR	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTGAACCA	TGCCAGCCGA	GAAATCGCCT	TCGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[563]
EUGLO	GCAAAGATGA	TCAGACACCG	TCGTAGTCCG	CTGTAAACGA	TGCCGGCCGA	GAAATCGCCT	GTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[590]
EVI	GCCAAGATGA	TCAGACACCG	TCGTAGTCCG	CTGTAAACCG	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCA	G	ATTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[599]
LSGc38	GCAAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTCAACTA	TGCCGGCCGA	GAAATCGCCT	GTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[563]
CPI	GCAAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[590]
CSK	GCAAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACCA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[590]
TET	GCGAAGATGA	TCAGATACCG	TCGTAGTCCA	CCGTAAACAA	TGTCGGCCGA	GAAATCGCCT	GTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[590]
EGN	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACGA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCA	G	CGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
TVS	GCGAAGATGA	TCAGACACCG	TCGTAGTCCR	CTGTAAACCA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCA	G	TGTAGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
SNB	GCGAAGATGA	TCAGACACCA	TCGTAGTCCG	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCCG	G	GGTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
EHS	GCAAAGATGA	TCAGACACCG	TCGTAGTCCG	CTGTAAACGA	TGCCGGCCGA	GAAATCGCCT	GTGGG-TCAG	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[599]
GRC	GCAAAGATGA	TCAGACACCG	TCGTAGTCCG	CTGTAAACGA	TGCCGGCCGA	GAAATCGCCT	GTGGGTTTCA	G	TGTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[590]
STB	GCGAAGATGA	TTAGACACCA	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCCG	G	GGTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
MUC	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCCG	G	GGTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[599]
CVL	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCCG	G	GGTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[563]
SOA	GCGAAGATGA	TCAGACACCA	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCCG	G	GGTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
SBR	GCGAAGATGA	TCAGACACCA	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	GTGGGTTTCCG	G	GGTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
CEL	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCCG	G	GGTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]
CUA	GCGAAGATGA	TCAGACACCG	TCGTAGTCCA	CTGTAAACTA	TGCCAGCCGA	GAAATCGCCT	TTGGGTTTCCG	G	GGTTCGCAAGG	CTGAAACTTA	AAGGAATTGA	CGGAATGGCA	[600]

	610	620	630	640	650	660	670	680	690	700	710	720	
ACACU	CCACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTTA	CCAGGTCGCG	ACAGTAAGGA	TTGACAGATT	GATAGCTCTT	TCTTGATTCT	ATGGGTGGTG	GTGCATGGCC	[716]
CAPOW	ACACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTCA	CCGGGTCGCG	ACAGTAAGGA	TTGACAGATT	GATAGCTCTT	TCTTGATTCT	ATGGGTGGTG	GTGCATGGCC	[715]
MONOV	CcACCaGgAg	TgGaGcTgCg	GcTtAaTtTg	AcTcAaCaCg	GgGaAaCtCa	CcAgGtCcAg	AcAgTaAgGa	TtGaCaGaTt	GaGaGcTcTt	TcTtGaTtCt	AtGgGtGgTg	GtGcAtGgCc	[716]
ORNAN	CCACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGAAACCTCA	CCCGGCCCGG	ACAGAAAGGA	TTGACAGATC	GATAGCTCTT	TCTCGATTCT	GTGGGTGGTG	GTGCATGGCC	[716]
NEUCR	CCACCAGGGG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTCA	CCAGGTCGAG	ACAAATGAGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATTCT	GTGGGTGGTG	GTGCATGGCC	[715]
CRYNE	CCACCAGGTG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTCA	CCAGGTCGAG	ACAGTGAGGA	TTGACAGATT	GATAGCTCTT	TCTTGATTCT	ATGGGTGGTG	GTGCATGGCC	[715]
HARSP	???????GAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTCA	CCAGGTCGAG	ACAGTAAGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATTCT	ATGGGTGGTG	GTGCATGGCC	[694]
ABSLG	CCACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTCA	CCAGGTCGAG	ACAATAAGGA	TTGACAGATT	AAAAGCTCTT	TCTAGATTTT	ATGGGTGGTG	GTGCATGGCC	[714]
HEXIN	CCACAAGACG	TGGAGTTGCG	GCTCAATTTG	ACTCAACACG	GGAAT-CTTA	CCAGACCCAG	ATG-TACGGA	GCAACAGATT	AAGAGTTCTT	TTGCGATTGG	GCAGTTGGTG	GTGCATGGCC	[703]
TRIPY	CCACCAGGTG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTTA	CCAGGTCGAG	ACAGGAAGGA	TTGACAGATT	GATAGCTCTT	TCTTGATTCT	GTGGGTGGTG	GTGCATGGCC	[716]
MALJA	-CACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTTA	CCAGGTCGAG	ACAGGAAGGA	TTGACAGATT	-TCAGCTCTT	TCTTGATTCT	ATGGGTGGTG	GTGCATGGCC	[711]
MALCA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
DINEX	CCACCAGATG	TGGAGCTGCG	GCTCAATTCG	ACTCAACACG	GGGAAACTTA	TCAGATTCGG	ACAGTGAGGA	TTGACAGATT	GATAGCTCTT	TCTTGATTCC	GTGGGTGGTG	GTGCATGGTC	[714]
MONSP	CCACCAGGTG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTTA	CCAGGTCGAG	ACAGGAAGGA	TTGACAGATT	GATAGTTCTT	TCTTGATTCC	GTGGGTGGTG	GTGCATGGTC	[719]
STRST	CCACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTCA	CCAGGTCGAG	ACAGGAAGGA	TTGACAGATT	GATAGCTCTT	TCTTGATTCC	ATGGGTGGTG	GTGCATGGTC	[716]
TRIVA	-CACCAGGGG	TGGAGCTGTG	GCTTAATTTG	AATCAACACG	GGGAAACTTA	CCAGGACCCAG	ATGTTTATGA	CTGACAGGCT	TCGGG-TCCT	TCAGGATATT	GTTTTTGGTG	GTGCATGGCC	[694]
SPIVO	CCACCAGACG	TGGAGTTGCG	GCTTAATTTG	ACTCAACGCG	CCAAT-CTTA	CTAGACCCGT	ATG- TACTGT	ACGTGAGATT	GATAGTTCTT	ACATGAATGA	GCAGGTGGTG	GTGCATGGCC	[687]
ANDIN	CCACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTTA	CCAGGTCGAG	ACAGGAAGGA	TTGACAGATT	GACAGCTCTT	TCTTGATTCT	ATGGGTGGTG	GTGCATGGCC	[702]
ANDGO	CCACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGAAACTTA	CCAGGTCGAG	ACAGGGAGGA	TTGACAGATT	GATAGCTCTT	TCTTGATTCT	ATGGGTGGTG	GTGCATGGCC	[716]
SECEC	CCACCAGGAG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGAAACTCA	CCAGGTCGAG	ACAGGGAGGA	TGGACAGATT	GAGAGCTCTT	TCTAGATTCT	ATGGGTGGTG	GTGCATGGCC	[717]
RECAM	CCACCAATGG	AGGAGCTGCG	GCTTAATTCG	ACTCAACTCG	GGAAACCTTA	CCAGGTCGAG	ACAGGGAGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATTCT	ATGGGTGGTG	GTGCATGGCC	[717]
HISAR	CGACAAGGCG	AGGAGCTGCG	GCTTAATTCG	ACTCAACTCG	GGGAAACTTA	CCAGAGCCAG	ACAGGGAGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATTCT	ATGGATGGTG	GTGCATGGCC	[717]
JAKLI	CCACCAGGAG	TGGAGCTGCG	GCTTAATTTG	ACTCAACACG	GGGCAACTTA	CCAGGTCGAG	ACGGGGAGGA	TGGACAGATT	GATAGCTCTT	TCTTGATTCT	GTGGGTGGTG	GTGCATGGCC	[717]
JAKBA	CCACCAGGAG	TGGAGCTGCG	GCTCAATTTG	ACTCAACACG	GGGCAACTTA	CCAGGTCGAG	ACAGGGAGGA	TGGACAGGCT	GACAGCTCTT	TCTAGATTCC	ATGGATGGTG	GTGCATGGCC	[596]
STALI	CCACCAGGAG	TGGAGTTGCG	GCTTAATTCG	ACCCAACACG	GGGAAACTCA	CGAAGGCAAG	ATGGATAGGA	TTGACAAGTT	-AAAGACCTT	TCTTGATTCT	ACAAGTGGGG	GTGCATGGCC	[352]
NAEGR	CCACCAGGAG	TGGAGTTGCG	GCTTAATTCG	ACTCAACACG	GGGAAACTCA	CCAGGTCGAG	ACAAGTTTGA	TTGACAGGTT	AATAGCCCTT	TCTTGATTGT	GTGGTGGGTA	GTGCATGGCC	[712]
NAEAN	CCACCAGGAG	TGGAGTTGCG	GCTTAATTCG	ACTCAACACG	GGGAAACTCA	CCAGGTCAGG	ACAAGTTTGA	TTGACAGGTT	AATAGCCCTT	TCTTGATTGT	GTGGTGGGTA	GTGCATGGCC	[713]
ACRRO	CCACCAGGAG	TGGAGTTGCG	GCTTAATTCG	ACTCAACGCG	GGAAACTCA	CCAGGTCAGG	ACAGTAATGA	TTGACAGGTT	GATAGCTCTT	TCTGATTCT	TTGGATGGTG	GTGCATGGCC	[712]
SAWMA	CCACCAGGAG	TGGAGTTGCG	GCTTAATTCG	ACTCAACGCG	TGAAAACTTA	CCAAATCCAA	ATTGTAAAGG	TTGACAGGTT	AATGACTTTT	-CTTGATTCA	TAAAGAAGTG	GTGCATGGTC	[702]
BST	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGTGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATCCC	TTGAATGGTG	GTGCATGGCC	[717]
RBS	CCACAAGGCG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGTGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	[706]
RMN	CCACAAGATG	TGAAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	AAAGTAAGGA	TTGACAGATT	GATTGTTCTT	TCTGATTCT	CTCAATGGTG	GTGCATGGCC	[717]
DMT	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGTAAGGA	TTGACAGATT	GATTGTTCTT	TCTCGATCTC	TTGAATGGTG	GTGCATGGCC	[718]
PERAM	CCACCAGACG	TGGAGTTGCG	GTTTAATTTG	ACTCAACACG	GGAACTTTA	CCAGGACAGG	ACAGGAATGA	TTGACAGATT	GATAGCTCTT	TCTTGATTCT	TTTGACGGTG	GTGCATGGCC	[662]
LEIMA	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[717]
LEIIN	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[701]
LEIBR	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[717]
LEITA	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[717]
LEPCO	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[717]
LEPBI	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[715]
LEPPO	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[717]
HERMU	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[716]
HERPE	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAATGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[716]
BLACU	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGGGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	[717]
TRYBR	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGTGAGGA	TTGACAGATG	GAGTGTTCTT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	[718]
TRYBE	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGTGAGGA	TTGACAGATG	GAGTGTTCTT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	[718]
TRYAV	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGTGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	[717]
TRYCR	CCACAAGACG	TGAAGCTGCG	GTTTAATTTA	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGTGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	[718]
TRYRA	CCACAAGACG	TGGAGCTGCG	GTTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGATCCGG	ACAGTGAGGA	TTGACAGATT	GAGTGTTCTT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	[718]
ENTO	CCACAAGCAG	TGGATTGCG	GCTTAATTTG	ACTCAACGCG	CAAAAACCTTA	CTTGGACAGA	ACACTGGTGC	ATGACGGATT	GA-TGATCCT	TCATAAGTGT	GTAAGGGGTG	GTGCATGGCC	[712]
PLoG	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGAAACCTCG	CCAGGTCGCG	ACTCTGAGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATCGT	GAGGGTGGTG	GTGCATGGCC	[715]
PERTR	CCACAAGGTG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGGTCAGG	ACGGTTAGGA	TTGACAGACT	GTGAGCTCTT	TCTTGATCCT	GTGGAGGGTG	GTGCATGGCC	[718]
ESL	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGGTCAGG	ACAGCTAGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	GTGGAGGGTG	GTGCATGGCC	[719]
EXS	CCACAAGGTG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGGTCAGG	ACGGTCTGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[720]
SAC	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGGTCAGG	ACAGCCAGGA	TGGACAGACT	GCTAGCTCTT	GCTTGATTCT	ATGGATGGTG	GTGCATGGCC	[683]
ECL	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGGTCAGG	ACGACCTGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATTCT	CGGGACGGTG	GTGCATGGCC	[720]
TSM	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACTTTA	CCAGGTCAGG	ACCGCAGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TGGGATGGTG	GTGCATGGCC	[720]

	610	620	630	640	650	660	670	680	690	700	710	720	
POR	CCACCAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGTCGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[710]
TVI	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGCCCGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TGGGATGGTG	GTGCATGGCC	[709]
MAA	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[718]
PSSc18	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACAGCCGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATTCT	GTGGATGGTG	GTGCATGGCC	[710]
EDS	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACAGCCGGGA	TGGACAGACT	GAGAGCTCTT	TCTTGATTCT	GTGGACGGTG	GTGCATGGCC	[718]
MOU	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAAGGTTA	CCAGGTCAGG	ACAGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATGCC	TTGGATGGTG	GTGCATGGCC	[720]
PYR	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAAGGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATGCT	TCGGATGGTG	GTGCATGGCC	[719]
LVO	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCCAGGA	TGGACAGATT	GAAAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[719]
TEA	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACAGTCAGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATTCA	GTGGATGGTG	GTGCATGGCC	[720]
EGR	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACAGCCGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	GTGGACGGTG	GTGCATGGCC	[719]
EPE	CCACCAAGGCG	TGGAGTTGCG	GCTTAATTTG	ATTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TTGACAGATT	GAAAGCCCTT	TCTTGATTCC	TTGGACGGTG	GTGCATGGCC	[719]
TLV	CCACMAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACSGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[720]
TPL	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[718]
LBU	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAAAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[708]
DAA	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGACT	GAGAGCTCTT	TCTTGATCCT	ATGGATGGTG	GTGCATGGCC	[719]
LAM	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[720]
LTR	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[683]
EUGLO	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACGCG	GGGAATGTTA	CCAGGTCAGG	ACGACTGGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATCCT	GTGGACGGTG	GTGCATGGCC	[710]
EVI	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	GTGGACGGTG	GTGCATGGCC	[719]
LSGc38	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[683]
CPI	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACGTTA	CCAGGTCAGG	ACGGCTGGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[710]
CSK	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAAAGTTA	CCAGGTCAGG	ACGGCTGGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[710]
TET	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATCCT	TTGGACGGTG	GTGCATGGCC	[710]
EGN	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAACGTTA	CCAGGTCAGG	ACAGCCGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	GTGGACGGTG	GTGCATGGCC	[720]
TVS	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATTCT	TCGGATGGTG	GTGCATGGCC	[720]
SNB	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACAGCCAGGA	TGGACAGACT	GCTAGCTCTT	GCTTGATTCT	ATGGATGGTG	GTGCATGGCC	[720]
EHS	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACGCG	GGGAATGTTA	CCAGGTCAGG	ACGACTGGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATTCT	GTGGACGGTG	GTGCATGGCC	[719]
GRC	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACGCG	GGGAATGTTA	CCAGGTCAGG	ACGACTGGGA	TTGACAGATT	GAGAGCTCTT	TCTTGATTCT	GTGGACGGTG	GTGCATGGCC	[710]
STB	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACAGTCAGGA	TGGACAGACT	GCTAGCTCTT	GCTTGATCCA	GTGGATGGTG	GTGCATGGCC	[720]
MUC	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGACT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[719]
CVL	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[683]
SOA	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACAGCCAGGA	TGGACAGACT	GCTAGCTCTT	GCTTGATCCA	GTGGATGGTG	GTGCATGGCC	[720]
SBR	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACAGCCAGGA	TGGACAGACT	GCTAGCTCTT	GCTTGATCCA	GTGGATGGTG	GTGCATGGCC	[720]
CEL	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[720]
CUA	CCACAAGGCG	TGGAGTTGCG	GCTTAATTTG	ACTCAACACG	GGGAATGTTA	CCAGGTCAGG	ACGGCTGGGA	TGGACAGATT	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	[720]

	730	740	750	760	770	780	790	800	810	820	830	840	
ACACU	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[836]
CAPOV	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGGGACG	GAGGGAGGCA	ATAACAGGTC	TGTGATGCCT	TCGATGTCCC	GGCGCACGCG	[835]
MONOV	GtTcTtAgTt	GgTgGaGtGa	TtTgTcGtTa	AtTcCgTtAa	CgAaCgAgAc	CtTaAcTtCt	TaGaGgGaCg	GaAgGaGgCa	AtAaCaGgTc	TgTgAtGcCt	TaGaTgTtCt	GgCgCaCgCg	[836]
ORNAN	GTTCTTAGTT	GGTGGAGCGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	TCTGACTTCT	TAGAGGGACG	GA- GTGAGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTCCG	GGTGCACGCG	[835]
NEUCR	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[835]
CRYNE	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[835]
HARSP	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[814]
ABSGL	GTTCTTAGTT	CGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGAGACG	GAAGTAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[834]
HEXIN	GCTCCTAGTT	CGTGGGTAA	ACTGTCCTT	ATTGCGTTAA	CGAGCGAGAC	AACCA- - - -	TGGGGGAGC	GAAYGTTGCT	CTAGCAGGTC	TGTGATGCCT	TAGACTCT	GGCGCACGCG	[818]
TRIPY	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTCAGCTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[836]
MALJA	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[831]
MALCA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
DINEX	GTTCTTAGTT	GGTGGGGTGG	CTTGTCTCTTA	ATTGCGATAA	CGAACGAGAC	CTCGGTTCT	TAGAGAGACG	GAAGGAGGCA	ATAACAGGTC	CGTTATGCCT	TAGACAATCT	GACGCACGCG	[834]
MONSP	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTCAACTTCT	TCGGGGGAGA	GTTGGAGATC	TTTTAATGGT	TAATATTTAA	TATATATTTT	GAT- - -AGCT	[836]
STRST	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTCAACTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTCCG	GGCGCACGCG	[836]
TRIVA	GTTGGTGGTG	CGTGGGTTGA	CCTGTCTGTTG	ATTGAGCTAA	CGAGCGAGAT	TATCGGTTCT	AATTGGGACG	GAAGGTAGCA	ATAACAGGTC	CGTGATGCTT	TAGATGCTCT	GGTGCACGCG	[814]
SPIVO	GCTCTTAGTT	CGTGGGTGA	ACTGTCTCTT	ATTGCGTTAA	CGAGCGAGAT	GGACT- - - -	CTGCGGTATT	GAGATCCGCG	CTAGCAGGTC	TGTGATGCCT	TAGACACTCT	AGCGCACGCG	[802]
ANDIN	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTCACTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[822]
ANDGO	GTTCTTAGTT	GGTGGATTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGTTCT	GGCGCACGCG	[836]
SECEC	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CGAAACTTCT	TAGAGAGATT	GAAATCGGCC	AAAACAGGTC	TGTGATGCCT	TCGATGTTCT	G- CACACGCG	[836]
RECAM	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTCTGCTTCT	TAGAGGGATT	GAAAAGGCTC	AGAACAGGTC	TGTGATGCCT	TCGATGTTCT	GGCGCACGCG	[837]
HISAR	GTTTTTAGTT	GGTGGAGTGA	TCTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CTCTGCTTCT	TAGAGGGATT	GAAAAGGCC	AGAACAGGTC	TGTGATGCCT	TCGATGTTCT	GGTGCACGCG	[837]
JAKLI	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTA	ATTCCGTTAA	CGAACGAGAC	CGAAACTTCT	TAGAGAGATT	GAAATCGGCC	AAAACAGGTC	TGTGATGCCT	TCGATGTTCT	GGCGCACGCG	[837]
JAKBA	GTTCTTAGTT	GGTGGAGTGA	TTTGTCTGTTG	ATTCCGTTAA	CGAACGAGAC	CTCCGCTTCT	TAGAGAGACG	GAAGGAGGCA	AAAACAGGTC	TGTGATGCCT	TCGATGTTCT	GGCGCACGCG	[716]
STALI	GTTCTTAGTT	CGTGGGTTAA	CTTGTCTGTTA	ATTCCGTTAA	CGAGCAAGAC	CTCGGCTTCT	TAGCAGGACG	GAAGGAGGCA	ATAACAGGTC	TGTGATGCCT	TAGATGCCTT	CGAGCACGCG	[472]
NAEGR	GTTTCCAGTT	CGTGGAGTGA	TCTGTCTGTTA	ATTGAGATAA	CGAACGAGAC	CTAAGCTTCT	TAAAGGGACG	GAAGTAGGCC	ATAACAGGTC	TGTGATGTTCT	TAGATGTCTT	GGTGCACGCG	[832]
NAEAN	GTTTCCAGTT	CGTGGAGTGA	TCTGTCTGTTA	ATTGAGATAA	CGAACGAGAC	CTAAGCTTCT	TAAAG- GACG	GAAGTAGGCC	ATAACAGGTC	TGTGATGTTCT	TAGATGTCTT	GGTGCACGCG	[832]
ACRRO	GTTCTTAGTT	CGTGGAGTGA	TTTGTCTGTTA	ATTGAGTTAA	CGAACGAGAC	CTTAACTTCT	TAGAGGAACG	GAAGAAGGCA	ATAACAGGTC	TGTAATGCC	TAGATGTCTT	GGAGCACGCG	[832]
SAWMA	GTTCTTAGTT	CGTGGATTAA	TCTGTCTCTTA	ATTGTGTTAA	CGAGCGAGAC	CTTAACTTCT	TAGTAAAAAG	GA- AAAGGCTC	ATAACAGATC	TGTTATGCTT	CAGATGTTTT	GG- GCACGCG	[820]
BST	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCAGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
RBS	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAATCTTCT	CTGCAGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGTACACGCG	[826]
RMN	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCAGGATT	GAAATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
DMT	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCAGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[838]
PERAM	GCTTTTGGTC	GGTGAAGTGA	TTTGTGTTTG	ATTCCGTTAA	CGGACGAGAC	CCAATCTTCT	TCGCGGATT	GGCATGGGCA	ACAGCAGGTC	CGTGATGTCC	TAAATGTCTT	GGGACACGCG	[782]
LEIMA	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAAATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
LEIIN	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAAATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[821]
LEIBR	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAAATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
LEITA	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAAATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
LEPCO	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAAATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
LEPBI	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAAATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[835]
LEPPO	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAAATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
HERMU	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[836]
HERPE	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[836]
BLACU	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
TRYBR	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[838]
TRYBE	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[838]
TRYAV	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[837]
TRYCR	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	CTGCGGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	CCAATGTTCT	GGGACACGCG	[838]
TRYRA	GCTTTTGGTC	GGTGGAGTGA	TTTGTGTTTG	ATTCCGTCAA	CGGACGAGAT	CCAAGCTTCT	TTGCGGGATT	GAGATGGGCA	ACAGCAGGTC	TGTGATGTCC	TCAATGTTCT	GGGACACGCG	[838]
ENTO	ACACGAAGTT	CGTGGGATGA	CCTGTCTCTTA	ATTGCGATAA	CGGGC- AGAC	CCTCACGTCT	AACCTACAGC	TC- AGGGACT	AATGTAGGTC	TGTGATGCC	CAGATGCCCA	AGCACACGCG	[830]
PLog	GTGCTGGGTT	GGTGGAGTGA	TTTGTCTCTTA	ATTGCGATAA	CGAACGAGAT	CACAAAAGCT	TAGAGGGACG	GAGAGGTGCT	TTACAGGTC	TGTGATGCCG	C- AATGTTCT	GGGACACGCG	[834]
PERTR	GTTCTTAGTT	CGTGGATTGG	TCTGTCTGTTA	ATTCCGTTAA	CGAACGAGAA	CTCTACTTCT	GCAAGGGACG	GAGGGAGTCA	ACAGCAGGTC	TGTGATGTCC	TAAATGTCTT	GGTGCACGCG	[838]
ESL	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGTTG	ATTCCGTTAA	CGAGTGAGAC	ATCTGCTTCT	CTGAGGTGCG	CATGGAGCCA	ACAGCAGGTC	TGTGATGTCC	CAGATGTCTT	GGCGCACGCG	[839]
EXS	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGTTG	ATTCCGTTAA	CGAGTGAGAT	ATCTATTCTT	CTGAGGCGCG	CATGGAGCCA	ACAGCAGGTC	TGTGATGTCC	CAGATGTCTT	GGCGCACGCG	[840]
SAC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGTTG	ATTCCGTTAA	CGAGTGAGAT	CTCTGCTTCT	CTGAGGTGCG	CATGGAGCCA	ACAGCAGGTC	TGTGATGTCC	TAGATGTTCT	GGCGCACGCG	[803]
ECL	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGTTG	ATTCCGTTAA	CGAGTGAGAC	ATCTGCTTCT	CTGAGGTGCG	CATGGAGCCA	ACAGCAGGTC	TGTGATGTCC	CAGATGTCTT	GGCGCACGCG	[840]
TSM	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGTTG	ATTCCGTTAA	CGAGTGAGAT	ATCTGCTTCT	CTGAGGTGCG	CATGGAGCCA	ACAGCAGGTC	TGTGATGTCC	CAGATGTCTT	GGCGCACGCG	[840]

	850	860	870	880	890	900	910	920	930	940	950	960	
ACACU	CGCTACACTG	ATTAATTCAA	CGTGCTGGGG	ATAGATCATT	GTAATTATTG	AAACGAGGAA	TTCTAGTAA	GCGCGAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[956]
CAPOV	CGCTACACTG	ACGAAGGCAG	AGTGATGGGG	ATTGATCCTT	GTAATTCTGG	AAACGAGGAA	TTCTAGTAC	CGCAAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	ACTACCGATG	[955]
MONOV	CgCtAcAcTg	AcGaAgGcAg	CgTgAtGgGg	AtAgAtCtTt	GtAaTtAtCg	AaAcGaGgAa	TtCcTaGtTa	GcGcGgTtCa	TcCtGcCcTt	TgTaCaCaCc	GcCcGtCgCt	AcTaCcGaTg	[956]
ORNAN	CGCTACACTG	ACTGGCTCAG	CGTGATGGGG	CGCAATTATTC	GCAATTATTG	CAACGAGGAA	TTCCAGTAA	TGCGGGGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	ACTACCGATG	[955]
NEUCR	CGCTACACTG	ACACAGCCAG	CGTGCTGGGG	ATAGAGCATT	GCAATTATTG	CAACGAGGAA	TCCTAGTAA	GCGCAAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	ACTACCGATG	[955]
CRYNE	CGCTACACTG	ACTGAGCCAG	CGTGCTGGGG	ATAGAGCATT	GCAATTATTG	CAACGAGGAA	TACCTAGTAA	GCGTAGTCA	CCCTGCCCTT	TGTACACACC	GCCCGTCGCT	ACTACCGATG	[955]
HARSP	CGCTACACTG	ATGAAGTCAG	CGTGCTGGGG	ATAGAGCATT	GCAATTATTG	CAACGAGGAA	TTCTAGTAA	GCGCGAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	ACTACCGATG	[934]
ABSLG	CGCTACACTG	ACAAAACCAG	CGTGCTGGGG	ATAGAACATT	GCAATTATTG	TAACGAGGAA	TTCTAGTAA	GCGCAAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	ACTACCGATG	[954]
HEXIN	TACTACAATG	TCGGCTRCAC	AGTGATGGGG	ATCGGAGATT	GGAATTTTTT	TAACGAGGAA	TGCTTGTAG	GTCTGCGTTA	TCCTGCTTTT	TGTACACACC	GCCCGTCGCT	CCTACCGACG	[938]
TRIPY	CGCTACACTG	ATGTCGTCAG	CATGATGGGG	ATTGATCTTT	GCAATTATGG	AAACGAGGAA	TACCTAGTAA	GCGCAAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[956]
MALJA	CGCTACACTG	ATGAGCTCAA	CATGATGGGG	ATAGAGCTTT	RCAATTATCC	GAACGAGGAA	TTCTAGTAA	GCGCGATTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[951]
MALCA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
DINEX	TACTACTATA	GCCGCCTCAA	CGTGCTGGGG	ATCGGTTTTT	GTAATTCTGA	AAACGCGGAA	TATCTCGTAA	GTGCAAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGA-G	[953]
MONSP	CATTAAAGCTT	TCAGGATGGA	TTTTATTGAA	AT-TATCATT	AAAAATTATTC	GGATTGTAAT	TCTCTTTTAG	TTAATTAAAT	TGTT- - - -GT	GGTGGAAGGC	TAATGGGGCT	CTGGATCTAA	[951]
STRST	CGCTACAATG	ATGAGGTCAG	CATGCTGGGG	ATTGATCTTT	GTAATTATTG	AAACGAGGAA	TACCTAGTAA	GTGCGAGTCA	TGCTGCCCTT	TGTACACACC	GCCCGTCGCT	ACTACCGATA	[956]
TRIVA	TGCTACAATG	TTAGGATCAA	TATAGTTGGG	ATTGACGTTT	GTAATC-AGC	GAACGAGGAA	TCCCTTGTA	ATGTGTGTCA	ACCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[933]
SPIVO	TACTACACTG	TACGTGGCAG	CATGGTTGGG	ATTGCGGGTT	GGAAATCTCC	CAACGAGGAA	TGCTAGTAA	TCATGAGTCA	TCCTGGCTCT	TGTACACACC	GCCCGTCGCT	TCTACCGATG	[922]
ANDIN	TGCTACACTG	ACGGATCCAA	CATGATGGGG	ATAGACCATT	GTAAC- -TTG	GAACGAGGAA	TGCTGTAA	GTCTGAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTA-CGA-G	[938]
ANDGO	CGCTACACTG	ACGGATGCAG	CATGATGGGG	-AAGACCCCT	GCAATTATTG	GAACGAGGAA	TTCTTGTA	GCGCAAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[955]
SECEC	CGCTACACTG	ATGGACGCAA	CGTGCTGGGG	ATCGACCCCT	GCAATTGTGG	GAACGAGGAA	TTCTAGTAG	GCGCGGGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[956]
RECAM	CGCTACACTG	ACGGAGGCAA	CATGCTGGGG	ATAGACCATT	GCAATTATTG	GAACGAGGAA	TCGTTAGTAA	TCGCGTGTCA	TCCTGGAGTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[957]
HISAR	CGCTACACTG	ACGAACCCAA	CATGCTGGGG	ATAGACTATT	GCAATTATTA	GAACGAGGAA	TGGCTAGTAA	GCGCGTGTCA	TCCTGCTCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[957]
JAKLI	CGCTACACCG	GTGAGCTCAA	CGTGATTGGG	ATAGTCTATT	GCAATTTTTA	GAACGAGGAA	TTCTAGTAA	GCGCGAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[957]
JAKBA	CGCTACACTG	ATGAACGCAA	CGTGTTTGGG	ATAGACTATT	GCAACTATTG	GAACGAGGAA	TTCTAGTAA	GCGCGAGTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[836]
STALI	TACTACAATG	TTAAATAAAG	CATGACAGGG	ATTGGTCATT	GGAA- - -GTG	GAACGAGGAA	TTCTCGTAA	TCGCAAGTCA	TCCTG- - - -	-----	-----	-----	[554]
NAEGR	TACTACAATA	ACGGTACCAG	CGTGACAGGG	ATCGAGGATT	GGAAC- -ATC	CAACGAGGAA	TTCTAGTAA	GCGTGGTTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[950]
NAEAN	TACTACAATA	ACGGTACCAG	CGTGACAGGG	ATCGAGGATT	GGAAC- -ATC	CAACGAGGAA	TTCTAGTAA	GCGTGGTTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[950]
ACRRO	TACTACACTG	TTAATACTAA	CATGACAGGG	ATAGCTGTTT	GGAA- - -ATC	CAACGAGGAA	TTCTAGTAA	TCGCGATTCA	TCCTGCCCTT	TGTACACACC	GCCCGTCGCT	CCTACCGATG	[949]
SAWMA	TACTACATTG	ATAGTTTCAG	TAAATCAGGA	ATTAAGGATT	GTAA- -ATTCT	CAACGAGGAA	TTCTAGTAA	TAGTAAATTA	TCCTGAAATTT	TGTACACACC	GCCCGTCGCT	CCTACCAATG	[938]
BST	CACTACAATG	TCAGTGAGAA	CACAATTGGG	ACCTAGTATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[957]
RBS	CACTACAATG	TCAGCGAGAA	CACAATTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[946]
RMN	CACTACAATG	TCAGTGAGAA	CACAATTGGG	ACCGAGTATT	GCAATTATTG	GAACGAGGAA	TATCTCGTAG	GCGTAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[957]
DMT	CACTACAATG	TCAGTGAGAA	CACAATTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[958]
PERAM	CACTACAATG	ACAAAGTGAA	TGTGCTTGGG	ATCGAAGCTT	GCAATTATTC	TAACGAGGAA	TGCTAGTAG	ACGTGCTTCA	T- - - - -	-----	-----	-----	[863]
LEIMA	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[957]
LEIIN	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[941]
LEIBR	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[957]
LEITA	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[957]
LEPCO	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[957]
LEPBI	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[955]
LEPPO	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[957]
HERMU	CACTACAATG	TCAGTGAGAA	CAGTTTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[956]
HERPE	CACTACAATG	TCAGTGAGAA	CAGTTTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[956]
BLACU	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCATT	GTITCCGATG	[957]
TRYBR	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGTATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[958]
TRYBE	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGTATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[958]
TRYAV	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGTATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[957]
TRYCR	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGTATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[958]
TRYRA	CACTACAATG	TCAGTGAGAA	CACACTTGGG	ACCGAGTATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GTITCCGATG	[958]
ENTO	CGATACAATG	GAGTTGGCAG	CATGGAAGGG	ATTGCCGGCA	GTACA- CGCC	GAACGAGGAA	TTGCTTGTA	GTACCTGTGA	TCCTGCCAGT	TGTACACACC	GCCCGTCGTT	GCTATCGATT	[949]
PLog	CACTACAATG	GCAATAGTCAA	CGGACCGGGG	ATAGGTGATT	GCAATTATTG	GAACGAGGAA	TGCTCGTAG	GCGCAGCTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GCTCCGATG	[954]
PERTR	CACTACACTG	GCAAGGACAC	TTTGTCAGGG	ATAGATGTTT	GCAATTATCC	AAACGAGGAA	TACCTAGTAA	GCATAGGTCA	TCCTGCCACT	TGTACACACC	GCCCGTCGTT	GCTACTGATG	[958]
ESL	CACTACACTG	ATCAGAGTGAA	GGTGCTGGGG	ATAGATGATT	GCAACTGTCT	GAACGTGGAA	TGCTAGTAG	GCGTTGGTCA	CCCTGCCATT	TGTACACACC	GCCCGTCGTT	GCTACCGATG	[959]
EXS	CACTACACTG	ACTCAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCTAGTAG	GCGTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GCTACCGATG	[960]
SAC	CACTACATTG	ACACAGTGAA	TGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCTAGTAG	GCGTAGTCA	TCCTGCC- -	-----	-----	-----	[890]
ECL	CACTACACTG	GCCCGGTGAA	GGACSSSSG	GGGGGGGGGG	GGGGGGGGGG	GGTGGTGGGG	TGCCGC-CT	GCGGTGTCCC	TGAAGAGGGT	GGGAAATCCT	GCAAGTCCGT	TTGTGAGAGC	[959]
TSM	CACTACATTG	ATTCAAGTGAA	GGCGATGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCGTAGT	GCGCGTGTCA	TCCTGCCATT	TGTACACACC	GCCCGTCGTT	GCTACCGATG	[960]

	850	860	870	880	890	900	910	920	930	940	950	960	
POR	CACTACATTG	GTCTGGTGAA	TGCTACTGGGG	ATAGATGCTT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	ACGTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[950]
TVI	CACTACATTG	ATTCAGTGAA	GGCGCTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGTGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[949]
MAA	CACTACATTG	ATACAGTGAA	GGTCTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGTGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[958]
PSSc18	CACTACATTG	GTCTGTGAA	TGCTACTGGGG	ATAGCTGATT	GCAACTGTCT	GAACGTGGAA	TGCCTAGTAG	GCCTGTGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[950]
EDS	CACTACATTG	ATTCAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGTGGAA	TGCCTAGTAG	GCCTGTGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[958]
MOU	CACTACATTG	ACACAGCCAG	GGTCTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTGTAA	GCCTGTGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
PYR	CACTACATTG	ACACAGCCAG	GGTCTGGGG	ATAGATGCTT	GCAACTGTCT	GAACGAGGAA	TGCCTGTAA	GGGTGTGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[958]
LVO	CACTACATTG	ACCCAGTGAA	TATACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	ACCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[959]
TEA	CATTACATTG	ACACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGTGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
EGR	CACTACATTG	GTGCAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGTGGAA	TGCCTAGTAG	GCCTGTGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[959]
EPE	CACTACAATG	ATACAATGAA	AGTGTCCGGG	ATCGATGCTT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	ACGCACGTCA	TCCTGCCAAT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[959]
TLV	CACTACATTG	ACACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
TPL	CACTACATTG	ACACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[958]
LBU	CACTACATTG	GCCCAATGAA	TATACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GTGTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[948]
DAA	CACTACATTG	ACACAGTGAA	TGTAAGTGGG	ATAGACGTTT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[959]
LAM	CACTACATTG	GTCCAGTGAA	TGTAAGTGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
LTR	CACTACATTG	ATCTGGTGAA	TGTAAGTGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCC---	-----	-----	-----	[890]
EUGLO	CACTACATTG	TCACAGTGAG	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGTGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[950]
EVI	CACTACATTG	ACTCAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGTGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[959]
LSGc38	CACTACATTG	GCCCAGTGAA	TGTAAGTGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TACCTAGTAG	GCCTGGGTCA	TCCTGCC---	-----	-----	-----	[890]
CPI	CACTACATTG	ACACAGTGAA	GGTACTGGGG	ATAGACCCCTT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[950]
CSK	CACTACATTG	ACACAGTGAA	AGTACTGGGG	ATAGACCCCTT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[950]
TET	CACTACATTG	ACACAGTGAA	CGTGTGGGG	ATAGATGATT	GCAACTATTT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[950]
EGN	CACTACATTG	ACACAGTGAA	TGTAAGTGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TACCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
TVS	CACTACATTG	ACACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
SNB	CACTACATTG	ACACAGTGAA	TGTAAGTGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
EHS	CACTACATTG	TCACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGTGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[959]
GRC	CACTACATTG	TCACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGTGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[950]
STB	CACTACATTG	ACACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
MUC	CACTACATTG	ACACAGGGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[959]
CVL	CACTACATTG	ACACAGGGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCC---	-----	-----	-----	[890]
SOA	CACTACATTG	ACACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
SBR	CACTACATTG	ACACAGTGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
CEL	CACTACATTG	ACACAGGGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]
CUA	CACTACATTG	ACACAGGGAA	GGTACTGGGG	ATAGATGATT	GCAACTGTCT	GAACGAGGAA	TGCCTAGTAG	GCCTGGGTCA	TCCTGCCATT	TGTACACACC	GCCCCTCGTT	GCTACCGATG	[960]

	970	980	990	1000	
ACACU	AATGGTGAAA	TCCTCGGATG	AACCTAGAGG	AAGGAGAAAGT	CGTAACAA [1004]
CAPOW	AATGGTGAGA	TCTCCGGACA	AACCTAGAGG	AAGTAAAAAGT	CGTAACAA [1003]
MONOV	AaTgGtGaGg	TcTcGgGaCa	AaCtTaGaGg	AaGtAaAaGt	CgTaAcAa [1004]
ORNAN	GATGGTGAGG	TCCTCGGACG	AACCTAGAGG	AAGTAAAAAGT	CGTAACAA [1003]
NEUCR	AATGGTGAGG	CTTCCGGACA	AACCTAGAGG	AAGTAAAAAGT	CGTAACAA [1003]
CRYNE	AATGGTGAGA	TCTCCGGACA	AACCTAGAGG	AAGTAAAAAGT	CGTAACAA [1003]
HARSP	AATGGTGAGG	TCTTCGGACA	AACT-----	-----	[958]
ABSGI	AATGGTGAGC	ATATGGGACA	AACCTAGAGG	AAGTAAAAAGT	CGTAACAA [1002]
HEXIN	GATGGTGAGT	TGTCAGACG	AACCTGGAGG	AAGGAGAAAGT	CGTAACAA [986]
TRIPY	AATGATGAAA	TGTTTGGATA	AATCTAGAGG	AAGGAGAAAGT	CGTAACAA [1004]
MALJA	GATGATGAAG	TCTCCGGATA	AATCTAGAGG	AAGGAGAAAGT	CGTAACAA [999]
MALCA	-----	-----	-----	-----	[588]
DINEX	AATGGTGAAA	TGCTCGGATA	AATCTAGAGG	AAGGAGAAAGT	CGTAACAA [1001]
MONSP	TTCGACTTAG	AG--GGACTG	ATGCTTCTGG	GCCGCACGCG	CGCTACAC [997]
STRST	AATGATGAAA	TGTTCCGATG	AATCTAGAGG	AAGGAGAAAGT	CGTAACAA [1004]
TRIVA	GATGGTGAAA	TCACCGGATA	AATCTAGA--	-----	[961]
SPIVO	TGTGTTGAGT	TCTATGGACG	AAGCTAGAGG	AAGGAGAAAGT	CGTAACAA [970]
ANDIN	GATGGTGAAA	TCTTCGGAA--	-----	-----	[956]
ANDGO	GATGGTGAAC	TCTTCGGATA	AACCTAGAGG	AAGGAGAAAGT	CGTAACAA [1003]
SECEC	GATGGTGAAA	TCTCAGGATA	AACCTAGAGG	AAGGAGAAAGT	CGTAACAA [1004]
RECAM	GATGGTGAAA	TCTTCGGATG	AATCTAGAGG	AAGGAGAAAGT	CGTAACAA [1005]
HISAR	GATGGTGAAA	TCTCCGGATA	AATCTAGAGG	AAGGAGAAAGT	CGTAACAA [1005]
JAKLI	GATGGTGAAG	ATTCTGGATA	AATCTAGAGG	AAGGAGAAAGT	CGTAACAA [1005]
JAKBA	GATGGTGAAA	CCTCGGGATA	AATCTAGAGG	AAGGAGAAAC C-----	[877]
STALI	-----	-----	-----	-----	[554]
NAEGR	GACGATGAAC	CTGGCGGATA	AATCTGTAGG	AAGGAAAAAGT	CGTAACAA [998]
NAEAN	GACGATGAAC	CTGGCGGATA	AATCTGTAGG	AAGGAAAAAGT	CGTAACAA [998]
ACRRO	AATGGTGAAT	TTGACGGATG	AACCTGTAGG	AAGGAGAAAGT	CGTAACAA [997]
SAWMA	AATAATAAAA	TTGATGGATA	AATCTGTAGG	AAGGAGAAAGT	CGTAACAA [986]
BST	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGTAACAA [1005]
RBS	ATGGTACAGA	TGATCGGACG	ATATTAGAGG	AAGCAAAA--	----- [983]
RMN	ATGGTGCAGG	TGATCGGACG	ACATTAGAGG	AAGCAAAAAGT	CGTAACAA [1005]
DMT	ATGGTGCAGG	TGATCGGACG	ACATTAGA--	-----	[986]
PERAM	-----	-----	-----	-----	[863]
LEIMA	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAA----	----- [993]
LEIIN	ATGGTACAGG	TGATCGGACG	AT-----	-----	[963]
LEIBR	ATGGTACAGG	TGATCGGACG	ATATTA----	-----	[983]
LEITA	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGTAACAA [1005]
LEPCO	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAA----	----- [993]
LEPBI	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT C-----	[996]
LEPPO	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT C-----	[998]
HERMU	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT -----	[996]
HERPE	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAA----	----- [992]
BLACU	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT CGT-----	[1000]
TRYBR	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGTAACAA [1006]
TRYBE	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGTAACAA [1006]
TRYAV	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGAAACAA [1005]
TRYCR	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAG- -----	[997]
TRYRA	ATGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT C-----	[999]
ENTO	GCAGGAAAGC	TTTACCTGTG	ATTCTAGAGG	AAGCAAAAAGT	CGTAACAA [997]
PLoG	GCTGTAGAGG	CGTTGAGACA	ATATTAGAGA	CGATAAAAAGT	CGTAACAA [1002]
PERTR	GTATCAGAGG	CTGGAAGACA	ATGTTAGAAG	AAGCAAAAAGT	CGTAACAA [1006]
ESL	ATCGTAGAAC	CATCAGGATG	ATGCTAGAGG	AAGCAAAAAGT	CGTAACAA [1007]
EXS	GCATTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAAGT	CGTAACAA [1008]
SAC	-----	-----	-----	-----	[890]
ECL	CGTGTAGATG	ATTGCAACTT	GTGTTTGTAC	CACCGCCCGT	CGTTGCTA [1007]
TSM	ACAGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAAGT	CGTAACAA [1008]

	970	980	990	1000	
[	.	.	.	.	]
POR	G TAG TAG AGC	C ATC CCG GAC A	A TG TT AG AGG	A AG C A A A A - -	----- [988]
TVI	A CAG TAG AGC	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A - -	----- [987]
MAA	A AAG TAG AGC	C ATC CAG GAC A	A TG CT AG AGG	A AG C A A A A GT	CGTAACAA [1006]
PSSc18	G CAG TAG AGC	C ATC CAG GAC G	A TG TT AG AGG	A AG C A A A A - -	----- [988]
EDS	G C GGT AGA AC	C CT CCG GAT G	A TG TT AG AGG	A AG C A A A A - -	----- [996]
MOU	G TGACAGAGC	C ATC CCG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1008]
PYR	T TGACAGAGC	C ATC CCG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1006]
LVO	G CAG TAG AGC	C ATC CAG GAC G	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1007]
TEA	A GAGCAGAGT	C ATC CAG GAC A	A TG CT AG AGG	A AG C A A A A GT	CGTAACAA [1008]
EGR	A CAG TAGA AC	C ATC CAG GAT G	A TG CT AG AGG	A AG C A A A A - -	----- [997]
EPE	G TTGCAGAGA	C TT CAG GACT	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1007]
TLV	A TCGTAGAGC	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1008]
TPL	A TCGTAGAGC	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1006]
LBU	G TAG TAG AGC	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [996]
DAA	G CAG TAG AGC	C ATC CAG GAC A	A TG TT ANN GG	A AG C A A A A - -	----- [997]
LAM	G TGG TAG AGC	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1008]
LTR	-----	-----	-----	-----	----- [890]
EUGLO	G TGG TAG AGT	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A - -	----- [988]
EVI	A TAG TAGA AC	C ATC CAG GAT A	A TG TT AG - AG	A AG C A A A A GT	CGTAACAA [1006]
LSGc38	-----	-----	-----	-----	----- [890]
CPI	G CAG TAG AGT	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A - -	----- [988]
CSK	G CAG TAG AGT	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A - -	----- [988]
TET	A TAG TAG AGG	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A - -	----- [988]
EGN	G TAG TAGA AC	C ATC CAG GAT A	A TG TT AG - AG	A AG C A A A A GT	CGTAACAA [1007]
TVS	A CAG TAG AGC	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1008]
SNB	G CAG TAG AGC	C ATC CAG GAC A	A TG TT GA AGG	A AG C A A A A GT	CGTAACAA [1008]
EHS	G TGG TAG AGT	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1007]
GRC	G TGG TAG AGT	C ATCA - - - - -	-----	-----	----- [965]
STB	G TAG TAG AGC	C ATC CAG GAC A	A TG TT AA AGG	A AG C A A A A GT	CGTAACAA [1008]
MUC	G TGG TAG AGC	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1007]
CVL	-----	-----	-----	-----	----- [890]
SOA	G CAG TAG AGC	C ATC CAG GAC A	A TG TT AA AGG	A AG C A A A A GT	CGTAACAA [1008]
SBR	G CAG TAG AGC	C ATC CAG GAC A	A TG TT AA AGG	A AG C A A A A GT	CGTAACAA [1008]
CEL	G TGG TAG AGC	C ATC CAG GAC A	A TG TT AG AGG	A AG C A A A A GT	CGTAACAA [1008]
CUA	G TGG TAG AGC	C ATC CAG GAC A	A TG TT AG - AG	A AG C A A A A GT	CGTAACAA [1007]

Figure 2.2. The 18S-84 interleaved alignment. The OTUs are grouped according to supergroup, then to “group” that were related in previous literature. The OTUs are given an abbreviated short name (see Table 2.2) which is consistent throughout the pages of the alignment. The column positions, from 1-1008, are given in ten base increments at the top of each page. The number of residues from the original sequence is given in brackets to the right of each line corresponding to its respective OTU.

	10	20	30	40	50	60	70	80	90	100	110	120	
HEXIN	GGATTAAGCC	ATGCATGCC-	AACTGCGAAC	GGCTCATTAA	ATCAGTTGAA	ATGTTCTGTC	TAACCTCTCT	ACTAA-TC--	---AGGCACG	ATCAGCTTCT	GACGCATCAT	TAGTAGGAGA	[113]
TRIPY	G-ATTAAGCC	ATGCATGTCT	AACTGCGAAT	GGCTCATTAT	ATCAGTAATA	GTTTTTTTGT	AAATTCTCAGA	GCTAATACTC	-AT-TATCAT	TCAAATTTCT	GCCCTATCAA	CTGATGGTAG	[117]
MALJA	G-ATTAAGCC	ATGCAAGCCC	-TCTGCGAAT	GGCTCATTAA	ATCAGTTATA	GTTTTTTTGT	AAATTTCTAGG	GCTAATAACCC	-AT-AATCAT	TCAAATTTCT	GCCCTATCAA	CTGACGGTAR	[116]
MALCA	C-ATTAAGCC	ATGCAAGCCC	-TCTGCGAAT	GGCTCATTAA	ATCAGTTATA	GTTTTTTTGT	AAATTTCTAGG	GCTAATAACCC	-AT-AATCAT	TCAAATTTCT	GCCCTATCAA	CTGACGGTAA	[116]
DINEX	G-ATTAAGCC	ATGCAAGTCT	AACTGCGAAA	AGCTCCTTAA	AATAGTTATA	ATTTTTTTGT	AAATTTCTAGA	GCTAATACTT	-TTTCTTCAT	TCAAATTTCT	AACCTATCAA	CCGACGTTAA	[118]
MONSP	T-TTGAATAG	AAATATAATT	TAGAGCTAAT	ACATGCGAGA	ATTCATTTTT	CTTTTGTTTT	TAAAGATCCG	ACTAAGGGCC	-AT-TGTCAT	TGAAATTTCT	GACCTATCAA	CTGATGGTAG	[117]
STRST	G-ATTAAGCC	ATGCATGTGT	CACTGCGAAT	GGCTCATTAT	ATCAGTTATA	ATTTTTTTGT	AAATTTCTAGA	GCTAATACTC	-AT-TATCAT	TGAAATTTCT	GCCCTATCAA	CTGACTTAAG	[117]
TRIVA	TATTTAAGCC	ATGCAAGTGT	TACTGCGAAT	AGGCTATTAA	TACGCTCAGA	ATGTTTTTGGC	GAACTCTGGT	GCTAATACT-	---ACGCCAT	TGAGTGAAGT	GACCTATCAG	CTGTACTTAG	[115]
SPIVO	G-ACTAAGCC	ATGCAAGCC-	ATC-GCAAAC	GGCTCATTAT	ACGCGTCTTC	ATGTTGGCCC	CAACTCTCTT	GCCAATAC--	---AGGGCAT	TCG-TGATTCG	CACGCATCAC	CTGTAGGTAG	[111]
ANDIN	-----	-----TGTCT	CACTGCGAAT	GGCTCATTAA	ATCAGTGATA	ATCTTTTGT	AAACTGTAGA	GCTAATACTC	-AT-CATCAT	TGGAATCTCT	GCCCTATCAG	CTGATGGTAG	[103]
ANDGO	G-ATTAAGCC	ATGCATGTCT	ATCTGCGAAT	GGCTCATTAA	ACAAGCTATA	ATTTTTTTGT	AAACTGTAGA	GCTAATACTC	-AT-CATCAT	TCAAATTTCT	TCCCTATCAG	CTGACGGTAG	[117]
SECEC	AGATTAAGCC	ATGCATGTCT	ATCTGCGAAC	GGCTCATTAC	AACAGTTCAT	ATCTCTTGGC	GAAACGCAGA	GCTAATACTC	-ACGGACCAA	TGAGTCTCT	GGCCTATCAG	CTGTGCGTAC	[119]
RECAM	GGATTAAGCC	ATGCATGTCT	AACGCGAAC	GGCTCATTAC	AACAGTTATC	ATTTCTTGGT	GAAACACCAGA	GCTAATACTC	-AT-AACCAA	TGAGTCTCT	GCCCTATCAG	CTGTGCGTAC	[118]
HISAR	GGATTAAGCC	ATGCATGTCT	AACGCGAAC	GGCTCATTTC	AACAGTTCATC	ATTTCTTGGT	GAAACACCAGA	GCTAATACTC	-AT-AACCAA	TGAAATTTCT	GCCCTATCAG	CTGTGCGTTG	[118]
JAKLI	AGATTAAGCC	ATGCATGTCT	AACGCGAAC	GGCTCATTAC	AACAGTTATC	ATTTCTTGGT	CAACTCCAGA	GCTAATACTC	-ATGAACCAA	TCAAGTTTCT	GCCCTATCAG	CTGTGCGTAG	[119]
JAKBA	GGATTAAGCC	ATGCATGTGT	CACTGCGAAC	GGCTCATTAC	AACAGTTCATC	ATCTCTTGGT	CAACTCCAGA	GCTAATACTC	-ACGAACCAA	TGAGTCTCT	GCCCTATCAG	CTGACGGTAG	[119]
STALI	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
NAEGR	AGCCTAAGCC	ATGCAAAATGT	ATCTGTGGAA	GGCTCATTAT	AACAGTTATA	CTCC-GCCAC	TAACGTCAGC	GAT-ATACTT	-TTCAGGCAG	AGGAGTTTCT	TACCTATCAG	CTGTTGTTTG	[117]
NAEAN	AGCCTAAGCC	ATGCAAAATGT	ATCTGTGGAA	GGCTCATTAT	AACAGTTATA	CTCC-GCCAC	TAACGTCAGC	GATAATACTT	-TTCAGGCAG	AGGAGTTTCT	TACCTATCAG	CTGTTGTTTG	[118]
ACRRO	AGATTAAGCC	ATGCATATGT	ATCTGTGAAC	GGCTCATTAT	AACAGTTATA	CTCC-GCCAC	TAAATGTCAGC	GATAATACTT	-TTCAGGAAG	ACGAATTTCT	TATCTATCAG	CTGTTGTCTC	[118]
SAWMA	G-ATTAAGCC	ATGTAATATGT	AACGTTGGCAA	GGCTCAATAA	ACGATTTATA	AACAACGCTCT	TAAAAGCATC	GATAATAC--	---TTAGTCG	ATGATTTTAT	CATCTATCAG	TTATAAGTGT	[114]
BST	GGACTAAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CCGATGGCCG	[118]
RBS	GGACTAAGCC	ATGCACGCTC	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACG-CAA	GCTAATACTC	AATTTTTCGA	-CGAACAACT	GCCCTAT--	-----CG	[107]
RMN	GGACTAAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGCA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CCGATGGCCG	[118]
DMT	GGACTAAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGCA	ATCTCGCAAA	AAAACGCCGAG	GCTAATACTC	AATTTTTTGA	-CGAACAACT	GCCCTATCAG	CCGATGGCCG	[119]
PERAM	-----	-----	-----	-----	-----	-----	-AAACGCCAG	GCTAATACTC	AAAGTCCCGA	-CGAGTCTCT	GCCCTATCAG	CTGATGGTAC	[58]
LEIMA	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CTGATGGCCG	[118]
LEIIN	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAAC-ACT	GCCCTATCAA	TCGATGCC-C	[116]
LEIBR	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CTGATGGCCG	[118]
LEITA	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CTGATGGCCG	[118]
LEPCO	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CTGATGGCCG	[118]
LEPBI	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CT--TGCCCG	[116]
LEPPO	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CTGATGGCCG	[118]
HERMU	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAAG-CAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CTGATGGCCG	[117]
HERPE	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACG-CAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CTGATGGCCG	[117]
BLACU	GGACTAAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAG	AAAACGCCCTA	GCTAATACTC	AATTTTGTGA	-CGAACAACT	GCCCTATCAG	CTGATGGCCG	[119]
TRYBR	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AATCTACTGA	-CGAACAACT	GCCCTATCAG	CCGATGGCCG	[119]
TRYBE	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AATCTACTGA	-CGAACAACT	GCCCTATCAG	CCGATGGCCG	[119]
TRYAV	GGGCTTAGC-	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CCGATGGCCG	[117]
TRYCR	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTGC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CCGATGGCCG	[118]
TRYRA	GGACTTAGCC	ATGCATGCCT	CTCTGCGCAT	GGCTCATTAC	ATCAGACGTA	ATCTCGCAAA	AAAACGCCAA	GCTAATACTC	AAT-TTTTGA	-CGAACAACT	GCCCTATCAG	CCGATGGCCG	[118]
ENTO	GGACTAAGCC	ATGCATATGC	GTCTGTGGAC	GGCTCATTAT	TGCGATTGCT	ATACCTCAGA	AAACCC-ACC	GTTAATACTA	-AGAGAGGAT	TTGC-AGTCG	ACCCTATCAG	TTGAAGGCCAA	[117]
PLOg	G-ATTAAGCC	ATGCAAACT	CTCCGTGAAC	GGCTCATTAA	ATCAGCGATT	ATCTCCAGTT	CAACTGTATC	GCTAATAACAA	TGG-AACTGG	ACGAATTTCT	GGCTATCAG	CTGATGACTG	[118]
PERTR	GGACTAAGCC	ATGCAAGTGT	ATCTGCGAAT	GGCTCATTAC	AACAGCAACA	ATCCATTGAT	TAACCGGCTA	GTTAATACTC	-AGAAGTCAT	AGGAATCTCT	GCCCTATCAG	CTGATGGTGA	[119]
ESL	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTTGGCAT	CAACCTTGAG	GCTAATACTC	AGGTGATCGC	TGGAATTTCT	GACCTATCAG	CTGATTTGG	[120]
EXS	GGACTAAGCC	ATGCATATCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	GTCTCGCGAA	AAACCTGGTG	GCTAATACTC	AGGTGATCGC	ACGAACCTCT	GACCTATCAG	CTGACTGTGG	[120]
SAC	-----	-----TTAC	ATCAGCAGTC	ATCTAGTGT	GAACCTAGTG	GCTAATACTC	AGGTGATCGC	TGGAATTTCT	GCGAACCTCT	GACCTATCAG	CTGACTGTGG	[84]	
ECL	TGACAGTGT	GTGCTGATT	GGCTCGGATT	TGATCGGAAG	AGCAGCGTT-	-CCGAGTGT	TAGTCCGCC	GCGAACCTCT	AGGTGATCAC	GCGAATTTCT	GACCTATCAG	CTGACTGTGG	[118]
TSM	GGACTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGT	AAACCTTGTG	GCTAATACTC	AGGTGATCAC	ACGAATTTCT	GACCTATCAG	CTGACTGTGG	[120]
POR	-----	ATGCAAGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGCAT	AAACCTTAGTG	GCTAATACTC	AGGTGATCGC	ACGAATTTCT	GACCTATCAG	CTGACTGTGG	[110]
TVI	-----	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGT	AAACCTTGTG	GCTAATACTC	AGGTGATCAC	ACGAATTTCT	GACCTATCAG	CTGACTGTGG	[110]
MAA	GGGCTAAGCC	ATGCACGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATGTCTGTG	AAACCTGGTG	GCTAATACTC	AGGTGATCAC	ACCAATTTCT	GACCTATCAG	CCGACTGTGG	[120]
PSSc18	-----	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCATGT	AAACCTTAGTG	GCTAATACTC	AGGTGATCAC	ACGAATTTCT	GACCTATCAG	CTGACTGTGG	[110]
EDS	GGGCTAAGCC	ATGCAAGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGCAT	AAACCTTGTG	GCTAATAACC	AGGTGATCGC	ACGAATTTCT	GACCTATCAG	CTGACTGTGG	[120]
MOU	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTTGTGT	AAACCTGGTG	GCTAATACTC	AGGTGATCAC	AAGAATTTCT	GACCTATCAG	CCGACTGTGG	[120]
PYR	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATTTTGTGT	AAACCTGGTG	GCTAATACTC	AGGTGATCGC	AAGAATTTCT	GACCTATCAG	CCGACTGTGG	[120]
LVO	GGACTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGT	AAACCTTAGTG	GCTAATACTC	AGGTGATCGC	ACGAGCTTCT	GACCTATCAG	CCGATGGTGG	[120]

	10	20	30	40	50	60	70	80	90	100	110	120	
TEA	GgAtTaAgCc	AtGcAtGtCt	CtCtGtGaAt	GgCtCcTtAc	AtCaGcAgTc	AtCtTgTgAt	AaAcCtTgTg	GcTaAtAcTc	AgGtGaTcAc	AtGaAcTtCt	GaCcTaTcAg	CtGaCtGtGg	[120]
EGR	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTGGTG	GCTAATACTC	GGGTGATCAC	ACGAACCTTCT	GACCTATCAG	CTGACTGTGG	[120]
EPE	GGACTAAGCC	ATGCATGTGT	CTCTGTGAAT	GGCTCATTAC	ATCAGCAACA	ATCGTACGGA	GAACCTGGCTG	GCTAATACTC	-AGTCTTGG	GCGAACTTCT	GACCTATCAG	CTGATTGTGG	[119]
TLV	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTGGTG	GCTAATACTC	AGGTGATCAC	ACGAACCTTCT	GACCTATCAG	CCGACTGTGG	[120]
TPL	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTGGTG	GCTAATACTC	AGGTGATCAC	ACGAACCTTCT	GACCTATCAG	CCGACTGTGG	[120]
LBU	-----	ATGCATGTGT	GTCTGTGAAT	GGCTCCTTAC	ATCAGCAGGC	ATCTTGTGAT	AAACCTAGTG	GCTAATACTC	-AGTGATCAC	ACGAACCTTCT	GACCTATCAG	CCGATGGTGG	[109]
DAA	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTTGTGGT	GAACCTAGTG	GCTAATACTC	AGGTGACCAC	ATGAACCTTCT	GACCTATCAG	CCGATTGTGG	[120]
LAM	GGACTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTAGTG	GCTAATACTC	AGGTGATCAC	ACGAGCTTCT	GACCTATCAG	CCGATGGTGG	[120]
LTR	-----	-----	-----	-----TTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTAGTG	GCTAATACTC	AGGTGATCAC	ACGAACCTTCT	GATCTATCAG	CCGATGGTGG	[84]
EUGLO	-----	ATGCACGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTTGTG	GCTAATACTC	AGGTGATCGC	AAGAGCTTCT	GACCTATCAG	CTGACTGTGG	[110]
EVI	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTTGTGAT	AAACCTTGTG	GCTAATACTC	AGGTGATCAC	ATGAACCTTCT	GACCTATCAG	CTGACTGTGG	[120]
LSGc38	-----	-----	-----	-----TTAC	ATCAGCAGTC	ATCTAGTGAT	AAACCTAGTG	GCTAATACTC	AGGTGATCAC	AAGAACCTTCT	GACCTATCAG	CCGATGGTGG	[84]
CPI	-----	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTTGTGAT	AAACCTGGTG	GCTAATACTC	AGGTGATCAC	AAGAACCTTCT	GCCCTATCAG	CCGACTGTGG	[110]
CSK	-----	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTTGTGAT	AAACCTGGTG	GCTAATACTC	AGGTGATCAC	AAGAACCTTCT	GCCCTATCAG	CCGACTGTGG	[110]
TET	-----	ATGCAAGCCT	CTCTGTGAAT	GGCTCATTAC	ATCAGCAATA	ATCTCCTGGT	GAACCTGGTG	GTTAATACTC	AGGTGACCAG	ACGAACCTTCT	GACCTATCAG	CTGACTGTGG	[110]
EGN	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCCGGTGAC	AAACCTTGTG	GCTAATACTC	AGGTGGTCAC	AGGAACCTTCT	GACCTATCAG	CTGATTGTGG	[120]
TVS	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTGGTG	GCTAATACTC	AGGTGATCAC	ACGAACCTTCT	GACCTATCAG	CCGATTGTGG	[120]
SNB	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTAGTGAT	GAACCTAGTG	GCTAATACTC	AAGTGATCAC	TCGAACCTTCT	GACCTATCAG	CTGACTGTGG	[120]
EHS	GGGCTAAGCC	ATGCACGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTTGTG	GCTAATACTC	AGGTGATCGC	AAGAGCTTCT	GACCTATCAG	CTGACTGTGG	[120]
GRC	-----	ATGCACGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTCGTGAT	AAACCTTGTG	GCTAATACTC	AGGTGATCGC	AAGAGCTTCT	GACCTATCAG	CTGACTGTGG	[110]
STB	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCTAGTGAT	GAACCTAGTG	GCTAATACTC	AGGTGATCAC	TCGAACCTTCT	GACCTATCAG	CTGACTGTGG	[120]
MUC	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCCAGTGGT	AAACCTGGTG	GCTAATACTC	AGGTGACCAC	AGGAACCTTCT	GACCTATNAG	CTGATTGTGG	[120]
CVL	-----	-----	-----	-----TTAC	ATCAGCAGTC	ATCCAGTGGT	AAACCTGGTG	GCTAATACTC	AGGTGACCAC	AGGAACCTTCT	GACCTATCAG	CTGATTGTGG	[84]
SOA	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCCAGTGAT	GAACCTAGTG	GCTAATACTC	AGGTGATCAC	TCGAACCTTCT	GACCTATCAG	CTGACTGTGG	[120]
SBR	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCCAGTGAT	GAACCTAGTG	GCTAATACTC	AGGTGATCAC	TCGAACCTTCT	GACCTATCAG	CTGACTGTGG	[120]
CEL	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCCAGTGGA	AAACCTGGTG	GCTAATACTC	AGGTGTCCAC	AGGAACCTTCT	GACCTATCAG	CTGATTGTGG	[120]
CUA	GGGCTAAGCC	ATGCATGTCT	CTCTGTGAAT	GGCTCCTTAC	ATCAGCAGTC	ATCCAGTGGT	AAACCTGGTG	GCTAATACTC	AGGTGACCAC	AGGAACCTTCT	GACCTATCAG	CTGATTGTGG	[120]

	130	140	150	160	170	180	190	200	210	220	230	240	
HEXIN	AGTAAAGGTT	ATCCTATGAG	TTACACGCGT	ACGGAGAATT	AGGGTTCGAC	TCCGGAGAAT	GAGCATGAGA	GACGGCTCAT	AGTTCTAAGG	GAGGCAGCAG	GC GCGGAAAT	TGCCCAATGT	[232]
TRIPY	GATAGAGGCC	TACCATGGTT	TTGACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAAA	GATGGCGACT	ACTTCCAAGG	AAGGCAGCAG	GC GCGCAAAAT	TACCCAATCC	[237]
MALJA	GATAGAGGCT	TACCGTGGTC	GTACACGGGTA	CGCGAGGATT	AGGGTCTGAC	TCCGGAGAGG	GAGCTTTAGA	AACGGCCGCC	ACTTCCAAGG	AAGGCAGCAG	GC GCGCAAAAT	TACCCAATTC	[236]
MALCA	GATAGAGGCT	TACCGTGGTC	GTACACGGGTA	CGCGAGGATT	AGGGTCTGAC	TCCGGAGAGG	GAGCTTTAGA	AACGGCCGCC	ACTTCCAAGG	AAGGCAGCAG	GC GCGCAAAAT	TACCCAATTC	[236]
DINEX	GGTAGTGTCT	TAGAGTAGTC	GTGACGGGTG	ACGGGGAATC	AGGGTTCGAT	ACC GGAGAGG	GC GCATGAGA	AATGGCGACC	ATCACTAAGG	TGAGCAGCAG	GC GCGCAAAAT	TACCCACTGG	[238]
MONSP	GATAGAGGCC	TACCATGGTA	TTGACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATGGCTACC	ACTTCCAAGG	AAGGCAGCAG	GC GCGCAAAAT	TACCCAATCC	[237]
STRST	GGTAGTTTCC	TTAAGTGGTA	TTGACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACAGCTCCC	ACTTCTAAGG	AAGGCAGCAG	GC GCGCAAAAT	TACCCAATCC	[237]
TRIVA	GGTCTTTACC	TAGGTAGGCT	ATCACGGGTA	ACGGGCGGTT	ACCGTCGGAC	TCCGGAGAAG	GC GCCTGAGA	GATAGCGACT	ATATCCACGG	GTAGCAGCAG	GC GCGAAACT	TTCCCACTCG	[235]
SPIVO	CGTAACGTGC	TTCCTAGGGA	TTAACGCCT-	ACAGGACATT	AGGGTTCGAC	TCTGGAGAGT	GAGCATGAGA	GACGGCTCAC	AGTTCTACGG	AAGGCAGCAG	GC GCGGAAAT	TGCCCAATGT	[230]
ANDIN	TGTCGTGGAC	TACCATGGCG	ATAACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATGGCTACC	ACATCCAAGG	AAGGCAGCAG	GC GCGTAAAT	TACCCAATCC	[223]
ANDGO	TGTAGTGGAC	TACCGTGGCG	GTAAACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTTAAA	AATGGCTACC	ACAACTAAGG	TTGGCAGCAG	GC GCGCAACT	TACTGACTCT	[237]
SECEC	GGTATTGGCC	TACCGAGGCG	TCGACGGGTA	ACGGGGAATT	AGGGTTTGAT	TCCGGAGAGG	GAGCCTGAGA	AATGGCTACC	ACATCCAAGG	AAGGCAGCAG	GC GCGCAAAAT	TGCCCAATCC	[239]
RECAM	GGTATTGGCG	TACCGAGGCC	GCAACGGGTA	ACGGGGAATT	GGGGTTTGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACATCTAAGG	AAGGCAGCAG	GC GCGCAAAAT	TACCCAATGG	[238]
HISAR	GATATGCGCC	AACCGAGGCT	TCCACGGGTA	ACGGGGGATT	AGGGTTTGAT	TCCGGAGAGG	GGGCCTGAGA	GATGGCCACC	ACATCTAAGG	AAGGCAGCAG	GC GCGCAAAAT	TACCCAATGG	[238]
JAKLI	GGTATTGGCC	TACCGAGGCT	TCTACGGGTA	ACGGGGAATT	AGGGTTTGAT	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACATCTAAGG	AAGGCAGCAG	GC GCGCAAAAT	TGCCCAATCT	[239]
JAKBA	TGTATTGGAC	TACCGTGGCT	TCTACGGGTA	ACGGGGAATT	AGGGTTTGAT	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACATCCAAGG	AAGGCAGCAG	GC GCGCAAAAT	TACCCAATCC	[239]
STALI	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
NAEGR	TTTAAAGGAC	AAACCAGGCT	TTGACGGGT-	ACGGGGAATC	AGTGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATCGCTACC	ACATCTAAGG	ACGGCAGCAG	GC GCGCAAAAT	TACCCAATCT	[236]
NAEAN	TTTAAAGGAC	AAACCAGGCT	ATTACGGGT-	ACGGGGAATC	AGTGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATCGCTACC	ACATCTAAGG	ACGGCAGCAG	GC GCGCAAAAT	TACCCAATCT	[237]
ACRRO	AATAGGGGTG	AGACACGGCG	ATTACGGAT-	TCGGAGAATC	AGTGTTCGAT	TCCGGAGAGG	GAGCCTTAAA	AATGGCTACT	ACATCTAAGG	ACGGCAGCAG	GC GCGGAAAT	TGCCCAATCT	[237]
SAWMA	T--TAATTTA	CACACAGACA	ACCACAGAT-	ACGGATAATC	AGTGTTCGAT	TCCGGAGAAA	GAGCCTTAAA	AATGGCTACT	AGTCCAAGG	TTTGGCAACG	GC GCGGAAAT	TGCCCAATGC	[231]
BST	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[237]
RBS	TGTAGTGGAC	TGCCATGGCT	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATGGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGTAAAT	TGCTCAATGC	[226]
RNM	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATGGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGC	[237]
DMT	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATGGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGC	[238]
PERAM	TGTATTGGAC	TACCATGGCG	ATAACGGG-T	ACGGGGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGTAAAT	TGCCCAATGC	[177]
LEIMA	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[237]
LEIIN	TGTAATGGAC	TGCCAT--GCT	TTTACCGA--	ACGGGGGAAT	AAGG--TCGAT	TTCGGAAAGG	AGGCTGG-AA	AATTGCTT-C	ACCTCTACCG	AAGGGCAGCA	GC GCGCAAAAT	GGTCC-ATGT	[229]
LEIBR	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[237]
LEITA	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[237]
LEPCO	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[237]
LEPBI	TGTAGTGGAC	TGCCAGGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGC	[235]
LEPPO	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[237]
HERMU	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[236]
HERPE	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[236]
BLACU	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[238]
TRYBR	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[238]
TRYBE	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[238]
TRYAV	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[236]
TRYCR	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[237]
TRYRA	TGTAGTGGAC	TGCCATGGCG	TTGACGGG-A	CGCGGGGATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATAGCTACC	ACTTCTACGG	AGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGT	[237]
ENTO	GGTCTTTGCC	TGCTTTGACC	GTGACGGGTC	ACGGAGAATT	AGGGTTTGAT	TCCGGAGAGG	AAGCCTGAGA	CACGGCTACC	ATGTCCACGG	TCAGCAGCAG	GCACATAAAT	TGCCCCATAC	[237]
PLOg	TGCAGTGAAC	AGTCATGGTA	TTAACGAGT-	ACGGGGAATT	AGGGTTCGAC	TCCGGAGAGG	GC GCCTGAGA	AACGGCGACC	AGATCTAAGG	ATTGCAGCAG	GCACGCAAAAT	TGCCCAATGG	[237]
PERTR	GGTAGTGGCT	CACCATGGCA	TTGACGGGTA	ACAGAGAATT	AGGGTTTGAT	TCTGGAGAGG	GAGCCTGAGA	AATGGCTACC	ACTGCTATGG	CAGGCAGCAG	GC GCGCAAAAT	TGCCCAATAC	[239]
ESL	TGTAGTGGAC	CACAATGGCT	CTTACGAGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGC	[240]
EXS	CGCAAGGGCC	CACAGTGGCG	GCAACGGGTG	ACGGAGAATT	CGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTGCC	ACTACCAAGG	TGGGCAGCAG	GC GCGCAAAAT	TGCCCAATAC	[240]
SAC	TGTAGTGGAC	CACAGTGGCC	ATGACGGGTA	ACGGAGAATT	TGGGTTCGAC	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GCGCAAAAT	TGCCCAATAC	[204]
ECL	TGTAGCGGAC	ACAAGTGGCA	GTACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	CACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GCGCAAAAT	TGCCCAATCC	[238]
TSM	TGTAGCGGAC	CACGGTGGCT	ATGACGGGTA	ACGGAGAATT	TGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTGCTAAGG	CGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGC	[240]
POR	TGTAGTGGAC	CACAGTGGCT	ATGACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACTGCTAAGG	CGGGCAGCAG	GCACGCAAAAT	TGCCCAATGC	[230]
TVI	TGTAGCGGAC	CACAGTGGCT	TTGACGGGTA	ACGGAGAATT	TGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	CACGGCTACC	ACTACCAAGG	CGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGC	[230]
MAA	TGTAGTGGAC	CACAGTGGCG	GTGACGGGTA	ACGGAGAATT	CGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GCACGCAAAAT	TGCCCAATG-	[239]
PSSc18	TGTAGTGGAC	CACAGTGGCA	GTACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGC	[230]
EDS	TGTCGTGGAC	CACAGTGGCC	TTGACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GCGCAAAAT	TGCCCAATGC	[240]
MOU	TGTAGTGGAC	CACAGTGGCA	GTAAACGGGTA	ACGGAGAATT	TGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GCGCAAAAT	TGCCCAATAC	[240]
PYR	TGTAGTGGAC	CACAGTGGCC	GTAAACGGGTA	ACGGAGAATT	TGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GCGCAAAAT	TGCCCAATAC	[240]
LVO	TGTAGAGCCC	-ACCATGGCG	GTGACGGGTA	ACGGAGAATT	AGGGTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTTCTAAGG	AGGGCAGCAG	GCACGCAAAAT	TGCCCAATGC	[239]

	130	140	150	160	170	180	190	200	210	220	230	240	
TEA	TgCaGtGgAc	CaCaGtGgCc	TtGaCgGgTa	AcGgAgAaTc	AgGgTtCgAt	TcCgGaGaGg	GaGcCtGaGa	AaCgGcTaCc	AcTaCcAaGg	TgGgCaGcAg	GcGcGcAaAt	TgCcCaAtAc	[240]
EGR	TGTAGTGGAC	CACAGTGGCG	GTGACGGGTA	ACGGAGAATC	AGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATCC	[240]
EPE	TGTCGTGGAC	TACAATGGCT	TCCACGGGTA	ACGGAGAATG	AGGGTTTTCGAT	ACCGGAGAGG	GAGCCTGAGA	GATGGCTACC	ACCGCTAAGG	CGGGCAGCAG	GC GC GCAAAAT	TGCCCTATAC	[239]
TLV	TGTAGTGGAC	CACAGTGGCA	GTGACGGGTG	ACGGAGAATT	TGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATAC	[240]
TPL	TGTAGTGGAC	CACAGTGGCA	GTGACGGGTG	ACGGAGAATT	TGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATAC	[240]
LBU	TGTCGAGGCC	-ACCATGGCA	GTGACGGGTG	ACGGAGAATC	AGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACTAAGG	TGGGCAGCAG	GCACGCAAAAT	TGCCCCATGC	[228]
DAA	TGTAGTGGAC	CACAATGGCC	TCAACGGGTC	ACGGAGAATC	AGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GATGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATGC	[240]
LAM	TGTATTGGAC	CACCATGGCA	GTGACGGGTA	ACGGAGAATC	AGGGTTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AACGGCTACC	ACTTCTAAGG	AGGGCAGCAG	GCACGCAAAAT	TGCCCAATGC	[240]
LTR	TGTAGTGGAC	CACCATGGCA	GTGACGGGTA	ACGGAGAATC	AGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GATGGCTACC	ACTTCTAAGG	AGGGCAGCAG	GCACGCAAAAT	TGCCCAATGC	[204]
EUGLO	TGTATTGGAC	CACAGTGGCC	TTGACGGGTA	ACGAAGAATC	AGGGTTTCGAT	TTCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GCACGCAAAAT	TGCCCCATGC	[230]
EVI	TGTAGTGGAC	CACAGTGGCT	ATGACGGGTG	ACGGAGAATT	AGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	AATGGCTACC	ATTACTAAGG	TGAGCAGCAG	GC GC GCAAAAT	TGCCCAATGC	[240]
LSGc38	TGTAGTGGAC	CACCATGGCG	GTGACGGGTA	ACGGAGAATC	AGGGTTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTTCTAAGG	AGGGCAGCAG	GCACGCAAAAT	TGCCCAATGC	[204]
CPI	TGTAGTGGAC	CACAGTGGCA	ATAACGGGTA	ACGGAGAATC	TGGGTTTCGAT	TCCGGAGAGG	CAGCCTGAGA	AACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATCC	[230]
CSK	TGTAGTGGAC	CACAGTGGCA	GTAAACGGGT	ACGGAGAATC	TGGGTTTCGAT	TCCGGAGAGG	CAGCCTGAGA	AACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATCC	[230]
TET	TGTAGTGGAC	CACCGTGGCT	ATAACGGGTA	ACGGAGAATC	AGGGTTTTCGAT	ACC GGAGAGG	GAGCCTGATA	AACGGCTACC	ACTGCTAAGG	TGGGCAGCAG	GCACGCAAAAT	TGCCCAATAC	[230]
EGN	TGTAGTGGAC	CACAATGGCT	TTGACGGGTA	ACGGAGAATC	AGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATAC	[240]
TVS	TGTAGTGGAC	CACAATGGCA	GTGACGGGTA	ACGGAGAATT	TGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATGC	[240]
SNB	TGTAGTGGAC	CACAGTGGCC	GTGACGGGTA	ACGGAGAATT	TGGGTTTCGAC	TCCGGAGAGG	CAGCCTGAGA	AACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATAC	[240]
EHS	TGTATTGGAC	CACAGTGGCC	TTGACGGGTA	ACGGAGAATC	AGGGTTTCGAT	TCCGGAGAGG	GAGTCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GCACGCAAAAT	TGCCCCATGC	[240]
GRC	TGTATTGGAC	CACAGTGGCC	TTGACGGGTA	ACGGAGAATC	AGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GCACGCAAAAT	TGCCCCATGC	[230]
STB	TGTAGTGGAC	CACAGTGGCC	ATGACGGGTA	ACGGAGAATT	TGGGTTTCGAC	TCCGGAGAGG	CAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATAC	[240]
MUC	TGTAGTGGAC	CACAATGGCA	TTGACGGGTA	ACGGAGAATT	TGGGTTTCGAT	TCCGGAGAGG	GAGCCTTAGA	GACGGCTAC -	CCTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATGC	[239]
CVL	TGTAGTGGAC	CACGATGGCA	TTGACGGGTA	ACGGAGAATT	TGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATGC	[204]
SOA	TGTAGTGGAC	CACAGTGGCA	GTCAACGGGT	ACGGAGAATT	TGGGTTTCGAC	TCCGGAGAGG	CAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATAC	[240]
SBR	TGTAGTGGAC	CACAGTGGCA	ATCACGGGTA	ACGGAGAATT	TGGGTTTCGAC	TCCGGAGAGG	CAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATAC	[240]
CEL	TGTAGTGGAC	CACAATGGCA	TTGACGGGTA	ACGGAGAATT	TGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATGC	[240]
CUA	TGTAGTGGAC	CACAATGGCA	TTGACGGGTA	ACGGAGAATT	TGGGTTTCGAT	TCCGGAGAGG	GAGCCTGAGA	GACGGCTACC	ACTACCAAGG	TGGGCAGCAG	GC GC GCAAAAT	TGCCCAATGC	[240]

	250	260	270	280	290	300	310	320	330	340	350	360	
HEXIN	ACGAGGCAGT	GAGAAAAAT	-CTTGTCCAA	GCT---CAC	GCAATTAGTG	GGGAGGAGT	GGGAGGAGT	GGGAGGAGT	GGGAGGAGT	GGGAGGAGT	GGGAGGAGT	GGGAGGAGT	[343]
TRIPY	GGGAGGTAGT	GACAAAAAAT	ATCAATGCAA	TTGCGGAGGT	ATCAATTAGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCTAA	AGCGTATACT	AAAGTTGCTG	CAGAAAAAGCT	CGTAGCTGGA	[357]
MALJA	GAGAGGTAGT	GAAGACAAAT	AACAATGTAA	TTGTCTCGAT	-GCAATTGGA	GGGAAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AAATTTGCTG	CAGAAAAAGCT	CGTAGTTGGA	[355]
MALCA	GAGAGGTAGT	GAAGACAAAT	AACAATGTAA	TTGTCTCGAT	-GCAATTGGA	GGGAAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AAATTTGCTG	CAGAAAAAGCT	CGTAGTTGGA	[355]
DINEX	CTGAGGTAGT	GACAAAAAAT	ACCGATATAA	TTGGGCGCGA	A-CCATT-AA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCTAA	AGTGTATACT	AACGTTGCTG	ATGAAAAAGCT	CGTAGTTGGA	[356]
MONSP	GGGAGGTAGT	GACAAAAAAT	CAAAGCGGTA	TTGGAAAGGA	AACAATTAGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCTAA	AGTGTATACT	AAAATTGCTG	CAGAAAAAGCT	CGTAGTTGGA	[357]
STRST	GGGAGGTAGT	GACAAAAAAT	ACCGATGCAA	TTGTGGGGGT	AACCATTAGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCTAA	AGTGTATACT	AAAATTGCTG	CAGAAAAAGCT	CGTAGTTGGA	[357]
TRIVA	CGGAGGAGGT	AATGACCAGT	-TCCATTGG-	TGGGTACCAA	ACCTAGCAGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCTGC	AGTTTGCTCC	CATATTGTTG	CAGAAACGCC	CGTAGTCT--	[351]
SPIVO	ACGAGGCAGT	GAGAAGCATT	TGGGCTGGTG	TGACCGTTAC	G-CAAGTCGT	GGGAAAGCAT	GGTGGTAATT	CCATCATGAC	GACAATCCCT	TACGTTGCTG	CGGCAACGCT	CGTAGCTG--	[347]
ANDIN	GGGAGGTAGT	GACCATAAAT	AACGACGCGA	TCGTCTCGAA	AGCAACTGGA	GGGCAAGTCT	GGTGGTAATA	CCAGCTCCAG	AGCGTATATT	AAAGTTGTTG	CAGAAAAAGCT	CGTAGTTGGA	[343]
ANDGO	AGGAAGTAGT	GACAATAAAT	AACGCCGCGA	GTGTATCGAA	AGCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAG	AGCGTATATT	AAAATTGTTG	CAGAAAAAGCT	CGTAGTTGGA	[357]
SECEC	GGGAGGCAGT	GACGAGAAAT	AACCTGTTGA	CGGTATCGGT	AGCTAATGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAG	AGCGTATACT	AAATCTGTTG	CACAAAAAGCT	CGTAGTCCGA	[359]
RECAM	CCGAGGTAGC	GACAAGAAAT	AACCTGTTGA	CGGCACCGGT	AACAATTGGA	GGGCAAGTCT	GGTGGTAACA	CCAGCTCCAA	AGCGTATATT	AAAGTTGTTG	CACAAAAAGCT	CGTAGTCAGA	[358]
HISAR	CCGAGGTAGC	GACAAGAAAT	ATCCTTTGAA	AGGCACCGGT	AACAATTGGA	GGGCAAGTCT	GGTGGTAACA	CCAGCTCCAA	AGCGTATATT	AACCTTTGCTG	CACAAAAAGCT	CGTAGTCAGA	[358]
JAKLI	AGGAGGTAGT	GACAAGAAAT	ACCTGTTGAA	CGGTATCGGT	ATCCATTGGA	GGGCAAGTCT	GGTGGTAACA	CCAGCTCCAA	AGCGTATATT	AACGTTGTTG	CACAAAAAGCT	CGTAGTCAGA	[359]
JAKBA	GGGAGGTAGC	GACAAGAAAT	AACCTGTACTA	CGGTATCG--	-----	-----	-----	-----	-----	-----	-----	-----	[277]
STALI	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
NAEGR	AGGAGGTAGT	GACAAGCTAT	AGTGACTCG-	GTACTTAGGT	AGCCATTGGA	GGAAAAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATACTGCTG	TAGAAACGCC	CGTAGTATAC	[355]
NAEAN	AGGAGGTAGT	GACAAGATAT	AGTGACTCG-	GTACTTAGGT	AGCCATTGGA	GGAAAAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATACTGCTG	TAGAAACGCC	CGTAGTATAC	[356]
ACRRO	AGGAGGCAGT	GACAAGAAAT	AACAGCTCG-	GTAAAT-TAAA	A-TGACTAGA	GGAAAGAGTCT	GGTGGTAATT	CCAGCTCTAG	AGCGTATATT	AATACTGTTG	TAGAAACGCC	CGTAGTTGCG	[354]
SAWMA	GGGAGGCAGT	GACGAGAAAT	AGCATCCCA-	GTACTCTATA	GTCTATAATA	AAATTTGTTG	GGT-GTAAAGT	TCT-----AA	AGTCAATTAG	--GTTTATCT	AGAAAAACTA	TTTTAATAGT	[342]
BST	ATGAGGCAGC	GAAAAAGAAAT	GGCCTCC-AG	ATAAATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AACGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[356]
RBS	ATGAGGCAGC	GAAAAAGAAAT	AGGTTTTTCAT	TTATATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AACGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[346]
RMN	ATGAGGCAGC	GAAAAAGAAAT	AGGTGCCCGC	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AACGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[357]
DMT	ATGAGGCAGC	GAAAAAGAAAT	AGGTGCCCGC	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AACGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[358]
PERAM	GGGAGGCAGC	GAAGAAAAAT	AGCGATCCAG	TTG--ACGGT	AACGATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGTTGCTG	CTGAAGGGTT	CGTAGTTGGA	[295]
LEIMA	ATGAGGCAGC	GAAAAAGAAAT	AGAGTTGTA-	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[356]
LEIIN	ATGAGGCCGC	GAA-AGAAAA	AAAG-TGTA-	TTTCATCCGG	-ACATTTGTA	GGAC--ATCT	TGT-GTAA-T	TCTGCTTCAA	AAGG-GTATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[340]
LEIBR	ATGAGGCAGC	GAAAAAGAAAT	AGAGTTGTA-	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[356]
LEITA	ATGAGGCAGC	GAAAAAGAAAT	AGAGTTGTA-	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[356]
LEPCO	ATGAGGCAGC	GAAAAAGAAAT	AGAGTTGTA-	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[356]
LEPBI	ATGAGGCAGC	GAAAAAGAAAT	AGAGTTGTA-	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[354]
LEPPO	ATGAGGCAGC	GAAAAAGAAAT	AGAGTTGTA-	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[356]
HERMU	ATGAGGCAGC	GAAAAAGAAAT	AGGTTGTA-	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[355]
HERPE	ATGAGGCAGC	GAAAAAGAAAT	AGAGTTGTA-	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[355]
BLACU	ATGAGGCAGC	GAAAAAGAAAT	AGGTTGTA-	TTTTATCGGT	AGCCATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[357]
TRYBR	ATGAGGCAGC	GAAAAAGAAAT	AGAGCCGAGT	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[358]
TRYBE	ATGAGGCAGC	GAAAAAGAAAT	AGAGCCGAGT	TTTTATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[358]
TRYAV	ATGAGGCAGC	GAAAAAGAAAT	AGAGCCGAGT	TTTAATCGGT	AACAATTGGA	GGACAAGTCT	GATGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[356]
TRYCR	ATGAGGCAGC	GAAAAAGAAAT	AGAGCCGAGT	TTTAATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[357]
TRYRA	ATGAGGCAGC	GAAAAAGAAAT	AGAGCCGAGT	TTTAATCGGT	AACAATTGGA	GGACAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATATT	AATGCTGTTG	CTGAAGGGTT	CGTAGTTGAA	[357]
ENTO	GTGAGGCAGC	GACGGAACAT	CGGGGCG-G-	GAACA--GGG	ATCTACTAGA	GAGCAAGTCT	GGTGGTAATT	CCAGCTCTGG	AGCGTATACT	TACGTTGCTG	CTCAGGGACC	CTCAGTTGGA	[353]
PLoG	TTGAGGCAGT	GACGAGAAAT	AACAACACGA	TTGTGATGGT	ACCCACTGGA	GGGCAAGTCT	GGTCGTTGTT	CCTAAACATA	CAAACGCTGT	GACACTAGCG	CTCCAAAAAA	CAAAAAATTGC	[357]
PERTR	GTGAGGCAGC	GACGAAAAAT	AGCATTTCGT	TTGTAAATGGT	AACAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAG	AGCGTATACT	AATATTGTTG	CTGAAAAAGCT	CGTAGTTGGA	[359]
ESL	GTGAGGCAGC	GACGATCCGC	AGCAATCCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATACT	ATCATTGCTG	CTGAAACACC	CTCAGTTGCG	[360]
EXS	GTGAGGCAGC	AACGAACAGC	AGCAATCCAC	TTGTGATGGT	ACCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAG	GGCGTATACT	AACATTGCTG	CGGAAACGTT	TGTAGTCGGC	[360]
SAC	GTGAGGCAGC	GACGAATGCG	AGCAATTCAC	TTGTGATGGT	AGCCACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CGGAAACACT	TGTAGTCTGC	[324]
ECL	GTGAGGCAGC	GACGAATGCG	AGCAATTCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[358]
TSM	GTGAGGCAGC	GACGAATGCG	AGCAATTCAC	TTGTGATGGT	AACGACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAGCACT	TGTAGTCTGC	[360]
POR	GTGAGGCAGC	GACGAATGCG	AGCAATTCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGTGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCGGA	[350]
TVI	-CGAGGCAGC	GACGAATGCG	AGCAATTCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AATATTGCTG	CTGAAACACT	TGTAGTCTGC	[349]
MAA	-CGAGGCAGC	AACGAACCGC	AGCAATTCAC	TTGTGATGGT	ACCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AATATTGCTG	CTGAAACACT	TGTAGTCTGC	[358]
PSSc18	GTGAGGCAGC	GACGAATGCG	AGCAATTCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGTGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCGGA	[350]
EDS	-CGAGGCAGC	GACGAATGCG	AGCAATTCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CAGAAACACT	TGTAGTCTGC	[359]
MOU	GTGAGGCAGC	AACGAACCGC	AACAATCTAT	CTGTGATGGT	ACCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[360]
PYR	GTGAGGCAGC	AACGAACCGC	AGCAATCTAT	CTGTGATGGT	ACCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[360]
LVO	GTGAGGCAGC	GACGAATGCG	AGCAATTCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACGCT	TGTAGTCTGA	[359]

	250	260	270	280	290	300	310	320	330	340	350	360	
TEA	GtGaGgCaGc	AaCgAaCtGc	AgCaAtCaAc	TtGcCaTgGt	AaCaAgTgGa	GgGcAaGtCt	GgTgGtAaTt	CcAgCtCcAt	GgCgTaTaCt	AaCaTtGcTg	CtGaAaCaCt	CgTaGtCtGc	[360]
EGR	GGGAGGCAGC	GACGAACCGC	AGCAATCCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACACC	TGTAGTCTGC	[360]
EPE	GTGAGGCAGC	AACGAGTCGT	GGCAGTTCAC	TCGCAATGGC	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	TGCGTATACT	AATGTTGCTG	CTGAAACGCT	TGTAGTCCGA	[359]
TLV	GTGAGGCAGC	GACGAACCTGC	AGCAATTCAC	TTGTGATGGT	AACAACCTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[360]
TPL	GTGAGGCAGC	GACGAACCTGC	AGCAATTSAC	TTGTGATGGT	AACAACCTGGA	GGGCAAGTCT	GGTGGTAATT	-CAGCTCCAA	GGCGTATACT	AACATTG-CG	CTGAAACACT	TGTAGTCTGC	[358]
LBU	GTGAGGCAGC	GACGAACCTGC	AGCAATTCAC	TTGTGATGGT	ATCGACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAT	TGCGTATACT	AACATTGCTG	TTGAAACACT	TGTAGTCTGA	[348]
DAA	-CGAGGCAGC	AACGAACCTGC	AGCAATGCAC	TTGTGATGGA	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATGCT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[359]
LAM	GTGAGGCAGC	GACGAACCGC	AGCAATTCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	CGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGA	[360]
LTR	GTGAGGCAGC	GACGAACCTGC	AGCAATTCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATACT	GATGTTGCTG	CTGAAACACT	TGTAGTCTGA	[324]
EUGLO	GTGAGGCAGC	GACGAACAGT	AGCAATCCAC	TTGTCATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GNCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[350]
EVI	-CGAGGCAGC	GACGAACCTGC	AGCAATCCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACACC	TGTAGTCTGC	[359]
LSGc38	GTGAGGCAGC	GACGAACCTGC	AGCAAGCCCC	TTGTGATGGT	ATCAACCGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCGA	TGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGA	[324]
CPI	GTGAGGCAGC	GACGAACCTGC	AGCAATCCAC	TTGTCATGGT	ACCTATTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[350]
CSK	GTGAGGCAGC	GACGAACCTGC	AGCAATTCAC	TTGTCATGGT	ACCTATTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[350]
TET	GTGAGGCAGC	AACGAGCTGT	AGCAATCCAT	TTGCAATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AATATTGCTG	CTGAAACGCT	TGTAGTCGAC	[350]
EGN	GTGAGGCAGC	GAAGAAGTGT	AGCAATCCAC	TTGTGATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	ATCATTGCTG	CTGAAACACC	TGTAGTCTGC	[360]
TVS	GCGAGGCAGC	AACGAACCTGC	AGCAATTCAC	TTGTGATGGT	AACGACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[360]
SNB	GTGAGGCAGC	GACGAACCTGC	AGCAATTCAC	TTGTGATGGT	AGCCACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CGGAAACACT	TGTAGTCTGC	[360]
EHS	GTGAGGCAGC	GACGAACAGT	AGCAATCCAC	TTGTCATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[360]
GRC	GTGAGGCAGC	GACGAACAGT	AGCAATCCAC	TTGTCATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	AGCGTATACT	AACATTGCTG	CTGAAACACT	TGTAGTCTGC	[350]
STB	GTGAGGCAGC	GACGAACCTGC	AGCAATTCAC	TTGTGATGGT	AGCCACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CGGAAACACT	TGTAGTCTGC	[360]
MUC	GTGAGGCAGC	GACGAACCGC	AACAATCCAC	TTGTCATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATGCT	AACATTGCTG	CGGAAACACT	TGTAGTCTGC	[359]
CVL	GTGAGGCAGC	GACGAACCGC	AACAATCCAC	TTGTCATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATGCT	AACATTGCTG	CGGAAACACT	TGTAGTCTGC	[324]
SOA	GTGAGGCAGC	GACGAACCTGC	AGCAATTCAC	TTGTGATGGT	AGCCACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CGGAAACACT	TGTAGTCTGC	[360]
SBR	GTGAGGCAGC	GACGAACCTGC	AGCAATTCAC	TTGTGATGGT	AGCCACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATACT	AACATTGCTG	CGGAAACACT	TGTAGTCCGC	[360]
CEL	GTGAGGCAGC	GACGAACCGC	AACAATCCAC	TTGTCATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATGCT	AACATTGCTG	CGGAAACACT	TGTAGTCTGC	[360]
CUA	GTGAGGCAGC	GACGAACCGC	AACAATCCAC	TTGTCATGGT	ATCAACTGGA	GGGCAAGTCT	GGTGGTAATT	CCAGCTCCAA	GGCGTATGCT	AACATTGCTG	CGGAAACACT	TGTAGTCTGC	[360]

	370	380	390	400	410	420	430	440	450	460	470	480	
HEXIN	-GAAATCGG	AGTGTATTAC	AAGAATGACA	TTGC--T---	CTGA-TAGTA	TTTGGTGGGG	ACAGGTGAAA	TAGGATGATC	CACCAAGGAC	TTTCAACAGC	GAAGGCGTAA	GCGCCGTCC-	[455]
TRIPY	TAGAAATTGG	AGTGTAAAGC	TGGAATAATA	AAATAGA---	GGCATTTGTA	TTGCGTGCCT	AGAGGTGAAA	TTCTTGGATT	GACGCAAGAC	AAACTACTGC	GAAGGCGCCA	AGGATGTTTT	[474]
MALJA	TGCAAAATAG	AGTGTAAAGC	TGGAATAACA	AAATA-G---	GGCATTTCGA	CCTTGGAGTC	AGAGGTGAAA	TTCTTAGATT	TCCGAGAGAC	GAACCTCTGC	GAAGGACCCC	AGGATGTTCT	[471]
MALCA	TGCAAAATAG	AGTGTAAAGC	TGGAATAACA	AAATG-----	GGCATTAGTA	CCTTGGAGTC	AGAGGTGAAA	TTCTTAGATT	TCCGAGAGAC	TAACCTCTGC	GAAGGACCCC	AGGATGTTTT	[470]
DINEX	TAGAAATGG	AGTGTAAAGC	TGGAATAATA	AAATAGA---	GGCATTAGTA	TTGCATTGTC	AGAGGTGAAA	TTCTTGGATT	TATGCAAGAC	TAAGTAGTGC	GAAGGCGCCA	AGGATGTTTT	[473]
MONSP	TAGAAATGG	AGTGTAAAGC	TGGAATAATA	AAATAGACAC	GGTATCTGTA	TTGCATAGTC	AGAGGTGAAA	TTCTTGGATT	TATGCAAGAC	AAACAAGTGC	GAAGGACCCA	AGGATGTTTT	[477]
STRST	TAGAAATGG	AGTGTAAAGC	TGGAATAATA	AAATAGA---	GGCATGGGTA	TTGCATAGTC	AGAGGTGAAA	TTCTTAGATT	TATGCAAGAC	CAACTAATGC	GAAGGCGCCA	AGGATGTTTT	[474]
TRIVA	--AGGA----	----CTA---	TATGGCTACA	TGA-----	GGTAGATCTA	TCTCATGACG	AGTGGTGGAA	TACTTTGACT	CATGAGAGAG	AAGCTGAGGC	GAAGGCACCT	AGAGGGTTTC	[451]
SPIVO	-GTAACCTCT	AACGCA----	T--GAGGGC-	-----TA--	CCCA-CTCTA	CTAGCGGAGG	ACAGGTGAAA	TAGCATGATT	CTCTAAGAG	GATCGGCTGC	GTAAGCGAGA	CT-GTGCTCT	[449]
ANDIN	TAGAAATCAG	AGTGCAAAGC	TGGGATAACC	ACCCTGA---	GGCATTCTGA	TTCCGAAGTC	AGAGGTGAAA	TTCTTGGATT	TTCCGAAGAC	GAACATATGC	GAAGGCGCCA	AGGATGTTTT	[460]
ANDGO	CAAGAATCAG	AGTGTAAAGC	TGGAATAACA	ACGAAGA---	GGCATTCTGA	TTCCGTAGTC	AGAGGTGAAA	TTCTTGGATT	TACGGAAGAC	GAACCTATGC	GAAGGCGCCA	AGGATGTTTT	[474]
SECEC	CGCAATCAG	AGTGCAAAGC	TGGAATAACG	AAGCGGA---	GGCATTCTGA	TTCCGGCGTC	AGAGGTGAAA	TTCTAGGATT	GCCGGAAGAC	GAACAACGTC	GAAGGCGGCC	AGGATGTTCC	[476]
RECAM	CGCAATCAG	AGTGCAAAGC	TGGAATGACA	ACGCGGA---	GGCATTCTGA	TTTCCATGTC	AGAGGTGAAA	TTCTAGGATT	ATGGAAGAC	AAACTGCTGC	GAAGGCGCCA	AGGATGTTCC	[475]
HISAR	TGCAATCAG	AGTGTCAAGC	TGGAATGACG	AAACGGA---	GGCATTCTGA	TTCCAACGTC	AGAGGTGAAA	TTCTAGGATT	GTTGGAAGAC	GAACGGCTGC	GAAGGCGCCA	AGGATGTTCC	[475]
JAKLI	CGCAATCAG	AGTGTCAAGC	TGGAATAACG	ATGCGGA---	GGCATTCTGA	TTCAAGGTGC	AGAGGTGAAA	TTCTAGGATT	ACTTAGAGAC	GAACCTGTC	GAAGGCGGCC	AGGATGTTCC	[476]
JAKBA	-----	-----	-----	-----	GGCATTCTGA	TTCCGGTGC	AGAGGTGAAA	TTCTAGGATT	ACCGBAAGAC	GAACGAATGC	GAAGGCGCCA	AGGATATTCC	[357]
STALI	-----	-----	-----	-GCTAGC-	GGCACAAGTA	TTCCGGTGAC	AGGGGTGAAA	TCCGTACTCC	CGGAAGAC--	TACCAGATGC	GAAGGCGGCC	ACGGCGTTTC	[84]
NAEGR	CATAAATCGT	TGTGCAAAGC	TGGGACTGCA	GAGTA-----	CCCCCGGCTA	CCATGAGGCT	AGAGGTGAAA	TTCTGATGAT	CTCTAGAGAC	CAACTAAGGC	GAAGGCGGCC	ACCACAAGCT	[470]
NAEAN	CGTAAATCGT	TGTGCAAAGC	TGGGACTGCA	GAGAA-----	CCCCCGGCTA	CCATGAGGCT	AGAGGTGAAA	TTCTGAAACC	CTCATGTGAC	CAACTAAGGC	GAAGGCGGCC	ACCACAAGCT	[471]
ACRRO	CACAATCAG	GGTGCAAAGC	TGGGATTGTC	CTATA-----	GGCTCTGAAA	CTATAGTGCT	AGAGGTGAAA	TTTAGATACC	GCTATATGTC	TATCTGATGC	GAAGGCGTCC	AGTACATTTT	[469]
SAWMA	TCTAAAGTAA	TGCGAAAAGC	TGGGATAACA	CATATAG---	TTTCTCAATC	CTAGTGAATC	AGGGGTGAAA	TCCGTATAGT	TACCAAAGGC	TACCAAAGGC	GAAGGCAAGC	AGGATGTTCC	[459]
BST	CCTAAAGAAG	CGCGAAAAGC	TGGGATAACA	AAGGACAACC	GGAAAACGTA	CTGGATCGTG	AGAGGTGAAA	TTCTTAGACC	GTTCCAAGAC	GAACCTACGC	GAAGGCTCCA	AGGATACCTT	[476]
RBS	TCAAAGAAG	GGTGAAAAGC	CGGGATAACA	AAGGA---CC	GGAGAACGTA	CTTGAGCGTG	AGAGGTGAAA	TTCTTAGACC	GCTCAAAGAC	GAACCTACGC	GAAGGCTCCA	AGGATACCTT	[463]
RMN	TCTAAAGAAG	TGCGAAAAGC	TGGGATAACA	AAGCACAGCC	GGAAAACGTA	CTTGGGCGTG	AGAGGTGAAA	TTCTTAGACC	GCCCAAAGAC	GAACCTACGC	GAAGGCTCCA	AGGATACCTT	[477]
DMT	TCTAAAAAAG	TGCGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAAAACGTA	CTTGAGTGTC	AGAGGTGAAA	TTCTTAGACC	ACTCAAAGAC	GAACCTACGC	GAAGGCTCCA	AGGATACCTT	[478]
PERAM	CATAAAGCAG	CGTGTAAAGC	TGGAATGAC-	--AGA--GCA	GAAAACCGTA	TTCCGGCGTG	AGAGGTGAAA	TTCTTAGACC	GCCCGGGGAC	GAACCTCAAG	GAAGGCTTCA	AGTATACCTT	[410]
LEIMA	CCTAAAGAAG	CGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGGGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[476]
LEIIN	CCTAAAGAAG	CGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGGGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[460]
LEIBR	CCTAAAGAAG	CGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGGGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[476]
LEITA	CCTAAAGAAG	CGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGGGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[476]
LEPCO	CCTAAAGAAG	TGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGGGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[476]
LEPBI	CCTAAAGAAG	TGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGGGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[474]
LEPPO	CCTAAAGAAG	TGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGGGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[476]
HERMU	TCTAAAGAAG	TGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGGGCGTG	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[475]
HERPE	TCTAAAGAAG	TGTGAAAAGC	TGGGATAACA	AAGGACAGCC	TGAGAACGTA	CTGGGGCGTG	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTCAA	AGGATACCTT	[475]
BLACU	CCTAAAGAAG	TGCGAAAAGC	TGGGATAACA	AGAGACAGCC	GGAGAACGTA	CTCGAGCGTC	AGAGGTGAAA	TTCTTAGACC	GATCGAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[477]
TRYBR	CCCAAAAAAG	TGCGAAAAGC	TGGGATAACG	AAGCACAGCC	GGAGAACGTA	CTGGTGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[478]
TRYBE	CCCAAAAAAG	TGCGAAAAGC	TGGGATAACG	AAGCACAGCC	GGAGAACGTA	CTGGTGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[478]
TRYAV	TCTAAAAAAG	TGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGTGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[476]
TRYCR	TCTAAAAAAG	TGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGTGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGTATACCTT	[477]
TRYRA	TCTAAAAAAG	TGTGAAAAGC	TGGGATAACA	AAGGACAGCC	GGAGAACGTA	CTGGTGCGTC	AGAGGTGAAA	TTCTTAGACC	GCACCAAGAC	GAACCTACGC	GAAGGCTTCA	AGGATACCTT	[477]
ENTO	TATGAAGCAG	AGCTTTTGGC	CGGCAAGCTG	ATTTCGAC---	TTTGTCCTGA	CCTTCAAGGA	AGCGGTGATA	TGCGATGATC	CTGAAGGGAC	AACCTAACGC	GAAGGCTTCA	TGAACGGCTT	[470]
PL0g	CACAAAACAG	CGTGTAAAGC	TGGGATGACA	GGGGATGACC	GAGATACGTA	TTGTGGCGCT	AGTGGTGAAA	TACTTAGACC	GTCACAAGAC	GAACATCTGC	GAAGGCTTCA	AGCATGCCCT	[477]
PERTR	TACAAATCAG	TGTGAAAAGC	TGAAATGACA	AGGGACGAC-	GTGGTGCTGA	TTTCTGCGTT	AGAGGTGAAA	TTCTTAGACC	GTAGAAGGAC	ATACTGCGGC	GAAGGCTTCC	AGTCTAGCTC	[478]
ESL	CAAAACTCAG	TGTGCAAAGC	TGAAATCTCC	AAGGACAGCC	GTGCAAGATG	CTGCAAGGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAT	CAGCGGNGGC	GAAGGCTTCA	AGTCTCGTT	[479]
EXS	CAAAACCCAG	TGTGCAAAGC	TGATATGTCG	TGCGACGACT	GTGCAAGGAA	CTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGGC	CGACAGCGGC	CGAGGCGGCA	AGGCGCGGTT	[480]
SAC	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGGAAGGTA	TTGGAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACTGCAGC	GAAGGCTGCA	AGAGCATGTT	[444]
ECL	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCTG	CAGCACGACC	TGGAAGGATA	TTGCAAGGCC	AGAGGTGAAA	TTCTTAGACC	GTGCAAGAT	CCACGGCAGC	GAAGGCTGCT	AGTGCCTGTT	[478]
TSM	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	CGGGAAGGTA	CTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACTGCAGC	GAAGGCTGCA	AGTGCATGTT	[480]
POR	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGAAGGATA	TGGCTGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCCAGAT	CAACTGCAGC	GAAGGCTGCA	AGTGCATGTT	[470]
TVI	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	CGGGAAGGTA	TGGCAAGGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCTGGAC	CAACTGCAGC	GAAGGCTGCA	AGTGCACGTT	[469]
MAA	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGCAAGATA	TTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAT	CCACAGCAGC	GAAGGCGGCA	AGGATGAGTC	[478]
PSSc18	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGTACGACC	GTGAAGGATA	TTGTAAGGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACTGCAGC	GAAGGCTGCA	AGTGCATGTC	[470]
EDS	CAAAACTCAG	TGTGCAAC-GC	TGGCATGTCA	AAGTACGACC	GTGGAAGGTA	CTACCGCGCC	AGAGGTGAAA	TTCTTAGACC	GCGGCAAGAC	CCACGGCGGC	GAAGGCTGCA	AGGCTGTGTT	[478]
MOU	CAAAACTCAG	TGTGCAAAGC	TGGAATGACA	AAGCACGCCC	GTGGAAGATA	TTGCAAGGCC	AGAGGTGAAA	TTCTTAGACC	GCGGCAAGAT	CCCTGCGGC	GAAGGCGACA	TGCTGGATT	[480]
PYR	CAAAACTCAG	TGTGCAAAGC	TGGAATGACA	AAGCACGCCC	GTGGAAGGTA	TTGCAAGGCC	AGAGGTGAAA	TTCTTAGACC	GTCGCAAGAC	CCACTGCGGC	GAA-GCCGCA	TGTTTGAGTC	[479]
LVO	CAAAACTCAG	TGTGCAAAGC	TGGCATGGCA	AAGCACGATC	TGGAAGATA	TTGCAAGGCC	AGAGGTGAAA	TTCTTAGATC	GCTGCAAGAT	GACCCGACAG	GAAGGCTGCA	AGTGCCTGTC	[479]

	370	380	390	400	410	420	430	440	450	460	470	480	
TEA	CtAaAcTcAg	TgTgCcAaGc	GtGaAtTgCa	GaGcAcGaTc	GtGgAaGaTa	CtGcAgCgCc	AgAgGtGaAa	TtCtTaGaCc	GcTgCaAgAt	CaAcTgCaGc	GaAgGcTgCa	AgAgCtCgTt	[480]
EGR	CAAAACCCAG	CGTGCAA-GC	TGGAATGTCA	GAGCACGACC	GTGAAAGATA	CTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAT	CGACCGCAGC	GAAGGCTGCA	AGCGCGTGTT	[479]
EPE	CACAACCCAA	TGTGAAAAGC	CGAGATGGCA	ATGGACGAC-	GTGCTGGATA	TTGACACGCC	AGAGGTGAAA	TTCTTAGACC	GCACTGGGAT	CGACTGCGGC	AAAGGCTGCA	CGATTCTGTC	[478]
TLV	CAAAACTCAG	TGTGCAAAGC	TGGAATGGCA	AAGCACGACC	GTGAAAGGTA	TTGCAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCCAGAG	CAACTGCAGC	GAAGGCTGCA	AGTGCATGTT	[480]
TPL	CAAAACTCAG	TGTGCAAAGC	TGGAATGGCA	AAGCACGACC	GTGAAAGGTA	TTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCCAGAC	CAACTGCAGC	GAAGGCTGCA	AGTGCATGTT	[478]
LBU	TAA-ACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACT	GTGAAAGGTA	TTGCAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACTGCAGC	GAAGGCTGCA	AGTGCAGGTT	[467]
DAA	CAAAACTCAG	TGTGCAAAGC	TGGAATGACG	AAGCACGACC	GTGAACGATA	TTGAAGCGCT	AGAGGTGAAA	TTCTCAGACC	GCTGCAAGAT	CCACTGCAGC	GAAGGCTGCA	AGTGCTTGTT	[479]
LAM	TAAGACTCAG	TGTGCAAGGC	TGGAATGTCG	CAGTACGACC	GTGAAAGATA	TTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAT	CAACTGCGGC	AAAGGCTGCA	AGTGCATGTC	[480]
LTR	TAGAACTCAG	TGTGCAAAGC	TGGAATGTCA	TAGCATGACC	GTGGAAGGTA	TTGCAACGCC	AGAGGTGAAA	TTCTCAGATC	GCTGCAAGAC	CAACTGCGGC	GAAGGCTGCA	AGTGCATGTT	[444]
EUGLO	CAAAACTCAG	TGTGCAAAGC	TGGAACGTCA	AGGCACGACC	GTGAAAGATA	TGGGAGCGCC	AGAGGTGAAA	TTCTTAGATC	GCTGCCAGAT	CCACTGCAGC	GAAGGCTGCA	AGTGCACGTC	[470]
EVI	CAAAACTCAG	TGTGCAAAGC	TGGAATTTCA	AAGCACGACC	GTGAAAGGTA	TTGCAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACTGCAGC	GAAGGCTGCA	AGTGCATGTT	[479]
LSGc38	TAAACTCAG	TGTGCCAAGC	TGGAATGTCA	AAGCACGACC	GTGGAAGGTA	TTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACGGCGGC	GAAGGCTGCA	GGTGCACGTT	[444]
CPI	TAAACTCAG	TGTGCAAAGC	TGGAACGTCA	CAGCATGACC	GTCAAAGATA	TTGCAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAGGAT	CAACAGCAGC	GAAGGCGGCA	AGTGCATGTT	[470]
CSK	CAAAACTCAG	TGTGCAAAGC	TGGAACGTCA	CAGCATGACC	GTCAAAGATA	TTGCAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAGGAT	CAACAGCAGC	GAAGGCGGCA	AGTGCATGTT	[470]
TET	CACAACCTAG	AGTGTAAAC	TGGGATGACA	AAGGACGACC	GTGATGGGTA	TTGACGCGCT	AGAGGTGAAA	TTCTTAGATC	GCCTTGGGAC	CTTCTGCAGC	GAAGGCTGCA	AGGGCATGTC	[470]
EGN	CAAAACTCAG	TGTGCAAAGC	TGGAACGTCA	AAGCACGACC	GTGAAAGGAA	CTGCAGCGCC	AGAGGTGAAA	TTCTTAGATC	GCTGCAAGAC	CCACTGCAGC	GAAGGCTGCA	AGGACGTGTT	[480]
TVS	CAAAACTCAG	TGTGCAACGC	TGGAATGGCA	AAGCACGACC	GTGAAAGGTA	TTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACTGCAGC	GAAGGCTGCA	AGTGCATGTT	[480]
SNB	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGGAAGGTA	TTGGAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACTGCAGC	GAAGGCTGCA	AGAGCATGTT	[480]
EHS	TAAACTCAG	TTTGCAAAGA	TGGAATGTCA	AGGCACGACC	GTGAAAGATA	TGGGAGCGCC	AGAGGTGAAA	TTCTTAGATC	GCTGCCAGAT	CCACTGCAGC	GAAGGCTGCA	AGTGCACGTC	[480]
GRC	TAAACTCAG	TTTGCAAAGA	TGGAATGTCA	AGGCACGACC	GTGAAAGATA	TGGGAGCGCC	AGAGGTGAAA	TTCTTAGATC	GCTGCCAGAT	CCACTGCAGC	GAAGGCTGCA	AGTGCACGTC	[470]
STB	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGGAAGGTA	TTGGAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAC	CAACTGCAGC	GAAGGCTGCA	AGGGCATGTT	[480]
MUC	CATAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGAAAGATA	TTGCAACGCT	AGAGGTGAAA	TTCTTAGACC	GTTGCAAGAT	CAACGGCAGC	GAAGGCTGCA	AGTGCCTGTT	[479]
CVL	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGAAAGATA	TTGCAACGCT	AGAGGTGAAA	TTCTTAGACC	GCTGCAAGAT	GAACGGCAGC	GAAGGCTGCA	AGTGCCTGTT	[444]
SOA	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGGAAGGTA	TGTGAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGACAGAC	CAACTGCAGC	GAAGGCTGCA	AGGGCATGTT	[480]
SBR	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGGAAGGTA	TGTGAGCGCC	AGAGGTGAAA	TTCTTAGACC	GCTGACAGAC	CAACTGCAGC	GAAGGCTGCA	AGGGCATGTT	[480]
CEL	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGAAAGATA	TTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GTTGCAAGAT	CAACGGCAGC	GAAGGCTGCA	AGTGCCTGTT	[480]
CUA	CAAAACTCAG	TGTGCAAAGC	TGGAATGTCA	AAGCACGACC	GTGAAAGGTA	TTGCAACGCC	AGAGGTGAAA	TTCTTAGACC	GTTGCAAGAC	CAACGGCAGC	GAAGGCTGCA	AGTGCCTGTT	[480]

	490	500	510	520	530	540	550	560	570	580	590	600	
HEXIN	AGYTGGTCAA	GAGCGAAAGT	GTGAGGATAG	ACGATGATTA	GAAACCGTTT	TATTTACCTT	TAAACGATAC	CATCAGGGAA	ACCGGTTTCTG	GGTTCTG6GG	GAAGTATGAT	TGCAAGATTG	[575]
TRIPY	CATTAATCAA	GAACGAAAGT	TAGGGGATCG	AAGGCGATTA	GAAACCGGCC	TAGTCCTCCA	TAAACGATGC	CAACAGAGAAA	ATCGTCTCTG	GGTTCTG6GG	GGAGTATGGT	CGCAAGGCTG	[594]
MALJA	CAYTAATCAA	GAACGAAAGT	TAGGGGATCG	AAGACGATCA	GATACCGTCG	TAGTCTTTCA	TAAACGATAC	CGACTGAGAA	ATCGTCTTTG	GGTTCTG6GG	GGAGTATGGT	CGCAAGGCTG	[591]
MALCA	CATTAATCAA	GAACGAAAGT	TAGGGGATCG	AAGACGATCA	GATACCGTCG	TAGTCTTTCA	TAAACGATAC	CGCCTGAGAA	ATCGTCTTTG	GGTTCTG6GG	GGAGTATGGT	CGCAAGGCTG	[590]
DINEX	CATTAATCAA	GAACGAAAGT	GAGGGGATCG	AAGACGATCA	GAAACCGTCG	TAGTCCTCCA	TAAACTATAC	CGTCAGGGAA	ACCGTTTTCTG	GGTTCTG6GG	GGAGTATGGC	CGCAAGACTG	[593]
MONSP	CATTAATCAA	GAACGAAAGT	TAGGGGATCG	AAGGCGATCA	GAAACCGCTC	TAGTCTTTCA	TAAACGATGC	CAACAGAGAA	ATCGTCTCTG	GGTTCTG6GG	GGAGTATGGT	CGCAAGCTCTG	[597]
STRST	CATTAATCAA	GAACGAAAGT	TAGGGGATCG	AAGGCGATTA	GAAACCGCTC	TAGTCTTTCA	TAAACGATGC	CAACTGAGAA	ATCGTCTCTG	GGTTCTG6GG	GGAGTATGGT	CGCAAGGCTG	[594]
TRIVA	TGTCGATCAA	GGGCGAGAGT	AGGAGTATCC	AACAGGATTA	GAGACCGTGG	TAGTTCCTCT	TAAACGATGC	CGACGGAGAA	ATCGTTTTCTG	GGCTCTG6GG	GAACACGAC	CGCAAGGCTG	[571]
SPIVO	TGCTGGTCAT	GAGCGAGAGC	GCAGGGCTGG	ACGACGATCA	GAGACCGTTC	TACTCTGCC	TAAACGATAC	TGCC - GAAGT	CCTGCTTTCTG	AGTTCTG6GG	GGAGTATGCT	CGCAAGAGTG	[568]
ANDIN	CATTAATCAA	GAACGAAAGT	TCGGGGATCG	AAGACGATCA	GATACCGTCG	TAGTCCGCCA	TAAACGATGC	CGACCGAGAA	ATCGTTTTTG	GGTTCCG6GG	GGAGTATGGT	CGCAAGGCTG	[580]
ANDGO	CATTAATCAA	GAACGAAAGT	TCGGGGATCG	AAGATGATTA	GATACCGTCG	TAGTCCGCCA	TAAACGATGC	CGACTGAGAA	ATCGTCTTTG	GGTTCCG6GG	GGAGTATGGT	CGCAAGGCTG	[594]
SECEC	CATTGATCAA	GAACGAAAGT	TAGGGGATCG	AAGACGATTA	GATACCGTCA	TAGTCTTAGG	TAAACGATGC	CGACTGAGAA	ACCGTCTCTG	GGTTGTG6GG	GGAGTATGGT	CGCAAGGCTG	[596]
RECAM	CGTTGATCAA	GAACGAAAGT	TAGGGGATCG	AAGACGATCA	GATACCGTCC	TAGTCTCTCG	TAAACGATGC	CGACCGGAA	ACCGTCTTTG	GGTTGTG6GG	GGAGTATGGT	CGCAAGGCTG	[595]
HISAR	CGTTGATCAA	GAACGAAAGT	TAGGGGATCG	AAGACGATTA	GATACCGTCC	TAGTCTCTCG	TAAACTATGC	CGACCGGAA	ACTGTCTTTG	GGTTGTG6GG	GGAGTATGGT	CGCAAGGCTG	[595]
JAKLI	CATTGATCAA	GAACGAAAGT	TAGGGGATCG	AAGACGATCA	GATACCGTCC	TAGTCTTTCA	TAAACGATGC	CGACTGAGAA	ACCGTCTTTG	GGTTGTG6GG	GGAGTATGGT	CGCAAGGCTG	[596]
JAKBA	CATTGATCAA	GAACGAAAGT	TAGGGGATCG	AAGACGATCA	GATACCGTCC	TAGTCTTTCA	TAAACGATGC	CGACAGGGAA	ACTGTCTTTG	GGTTGTG6GG	GGAGTATGGT	CGCAAGGCTG	[477]
STALI	CGTTGATCAA	GAACGAAAGT	CAAGGGATCG	AAGATGATTA	GATATCGTCG	TAGTCTTTCCG	TAAACTATAC	TGCCAGGGAA	ACCGTTTTCTG	AGTTCTG6GG	GAAGTATGGT	CGCAAGGCTG	[204]
NAEGR	CGTCTATCAG	GGACGAAAGT	TGGGGGATCG	AAGACGATTA	GATACCGTCG	TAGTCCCTCA	TAAACGATGC	CAACAGGGAA	ACCGTCTTTG	GGTTCTG6GG	GGAGTATAGT	CGCAAGGCTG	[590]
NAEAN	CGTCTATCAG	GGACGAAAGT	TGGGGGATCG	AAGACGATTA	GATACCGTCG	TAGTCCCTCA	TAAACGATGC	CAACAGGGAA	ACCGTCTTTG	GGTTCTG6GG	GGAGTATAGT	CGCAAGGCTG	[591]
ACRRO	AATCGATCAG	GAACGAAAGT	TAGGGGATCG	AAGATGATTA	GATACCGTCG	TAGTCTTTCTA	TAAACGATGC	CAACAGGGAA	ACCGTTTTTG	GGTTCTG6GG	GAGTTATAGT	CGCAAGGCTG	[589]
SAWMA	TATTGATCAG	GAGCTAGAGT	TAAGGGAACA	AAGATGATTA	GATACCGTCG	TAGTCTTTCTA	TAAACTATAC	TAATCGAGAA	ATCATTTTTG	AGTTCCAGGG	GAAGTATATT	CGCAAGGCTG	[579]
BST	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACCA	CAACCGGTGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGAT	CGCAAGGCTG	[596]
RBS	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACCA	CAACCGGGGA	CCCCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[583]
RMN	CCTCAATCAA	GAACCAAAGT	GTGGGGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACCA	CAACCGGTGA	CCCCCGCGAA	AGCGGTTACA	GTCTCAGGGG	- GAGTACGTT	CGCAAGGCTG	[596]
DMT	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACCA	CAACCGGTGA	CCCCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[598]
PERAM	CATCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GATACCATTTG	TAGTCCACTG	CAACAGATGA	CGCCCGCGAA	AGCGGCTATG	GTCTCATGGG	GAAGTACGTT	CGCAAGGCTG	[530]
LEIMA	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[596]
LEIIN	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[580]
LEIBR	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[596]
LEITA	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[596]
LEPCO	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[596]
LEPBI	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[594]
LEPPO	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[596]
HERMU	CCTCAATCAA	GAACCAAAGT	CGGGGGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCGCTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[595]
HERPE	CCTCAATCAA	GAACCAAAGT	CGGGGGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCGCTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[595]
BLACU	CCTCAATCAA	GAACCAAAGT	GTGGAGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[597]
TRYBR	CCTCAATCAA	GAACCAAAGT	GTGGGGATCA	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACCATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[598]
TRYBE	CCTCAATCAA	GAACCAAAGT	GTGGGGATCA	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACCATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[598]
TRYAV	CCTCAATCAA	GAACCAAAGT	GTGGGGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[596]
TRYCR	CCTCAATCAA	GAACCAAAGT	GTGGGGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[597]
TRYRA	CCTCAATCAA	GAACCAAAGT	GTGGGGATCG	AAGATGATTA	GAGACCATTTG	TAGTCCACTG	CAACAGATGA	CACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[597]
ENTO	AATCTATCAT	GAGTCTGAGC	ATGAGGATCG	AAAAGGATCA	GATACCCCTG	TAGTTCCGCC	TAAACGATAG	TGCCAGTGAA	AACGGTTCTG	CCTCGAGGGG	GAAGTATAGT	CGCAAGGCTG	[590]
PL0g	CTTCAATCAA	GAGCGATAGC	ATGGGGATCT	AAGATGATCA	GATACCGTTG	TAGTCCACTG	TAAACGCTGC	CAGCCGAGAA	ATCGCC - TG	GGTTCCG6GG	GGAGTACGTC	CGCAAGGCTG	[595]
PERTR	TGTCTATCAA	GAACGAAAGT	GTGGGGAGCA	AAGATGATCA	GATACCGTTG	TAGTCCACTG	TAAACTATGT	CA - GCGAGAA	ATCACCTATG	GATTACAGGGG	GGAGTACTGT	CGCAAGGCTG	[597]
ESL	TGTCGATCAA	GGATGAGAGT	CTGGGGAGCA	AAGAGGATCA	GACACCCCTG	TAGTCCACTG	TAAACCGTGC	TAGTCCGCTG	ATCGCCCTGTG	GGTTCCG6GG	GGAGTACTGT	CGCAAGGCTG	[599]
EXS	CATCGATCAA	GAGCGAGAGC	CTGGGGATCG	AAGAGGATCA	GACACCCCTG	TAGTCCACCG	TAAACCATGC	CGCCCGAGAA	ATCGCCTTTG	GGTTCCG6GG	GGAGTACGCT	CGCAAGGCTG	[600]
SAC	CGTCGATCAA	GAATGAGAGC	TCGGGGAGCG	AAGATGATCA	GACACCATCG	TAGTCCGCTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCCG6GG	GGAGTACGCT	CGCAAGGCTG	[564]
ECL	CGTCGATCAA	GAATGAGAGT	ATGGGGAGCA	AAGATGATCA	GACACCCCTG	TAGTCCACCG	TAAACCATGC	CGCCCGAGAA	ATCGCCTGTG	GGTTCCG6GG	GGAGTACGCT	CGCAAGGCTG	[598]
TSM	CGTCGATCAA	GAATGAGAGT	GTGGGGAGCG	AAGATGATCA	GACACCCCTG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCCG6GG	GGAGTACTGT	CGCAAGGCTG	[600]
POR	CGTCGATCAA	GAATGAGAGT	GTGGGGAGCG	AAGATGATCA	GACACCCCTG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCCG6GG	GGAGTACTGT	CGCAAGGCTG	[590]
TVI	CGTCGATCAA	GAATGAGAGT	GTGGGGAGCG	AAGATGATCA	GACACCCCTG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCCG6GG	GGAGTACTGT	CGCAAGGCTG	[589]
MAA	CGTCGATCAA	GAATGAGAGC	CTGGGGATCC	AAGATGATTA	GACACCCCTG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCCG6GG	GGAGTACGCT	CGCAAGGCTG	[598]
PSSc18	TGTCGATCAA	GAATGAGAGT	GTGGGGAGCG	AAGATGATCA	GACACCCCTG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCCG6GG	GGAGTACTGT	CGCAAGGCTG	[590]
EDS	TGTCGATCAA	GAATGAGAGT	CTGGGGAGCA	AAGATGATCA	GACACCCCTG	TAGTCCGCTG	TCAACCGTGC	CAGCCGAGAA	ATCGCCTATG	GGTTCTG6GG	GGAGTACTGT	CGCAAGGCTG	[598]
MOU	CGTCGATCAA	GAATGAGAGC	CTGGGGAGCA	AAGGTGATTA	GACACCTCCG	TAGTCCACCG	TAAACCATGC	CGCCCGAGAA	ATCGCCTTTG	GGTTCCG6GG	GGAGTACTGT	CGCAAGGCTG	[600]
PYR	CGTCGATCAA	GAATGAGAGC	CTGGGGAGCA	AAGGTGATTA	GACACCTCCG	TAGTCCACTG	TAAACCATGC	CGCCCGAGAA	ATCGCCTTTG	GGTTCCG6GG	GGAGTACTGT	CGCAAGGCTG	[599]
LVO	CGTCGATCAA	GAACGTGAGT	GTGGGGAGCG	AAGATGATCA	GACACCTCCG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCCG6GG	GGAGTACTGT	CGCAAGGCTG	[599]

	490	500	510	520	530	540	550	560	570	580	590	600	
TEA	CgTcGaTCAa	GaAtGaGaGt	GtGgGgAaCa	AaGaTgAtCa	GaCaCcGtCg	TaGtCcAcAg	TaAaCcAtGc	CaGcCgAgAa	AtCgCcTgTg	GgTtCaGgGg	GgAgTaCtGt	CgCaAgGcAg	[600]
EGR	CGTCGATCAA	GGATGAGAGT	CCGGGGAGCA	AAGATGATCA	GACACCGTCG	TAGTCCGCCG	TAAACCGTGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCAGGGG	GGAGTACGTT	CGCAAGGCTG	[599]
EPE	CGTTGATCAA	GAACGAGAGT	GTGGGGAGCA	AAGATGATCA	GATACCGTTG	TAGTCCACCG	TAAGCGATGT	GAGCCGAGAA	ATCGCTTGTG	GATTACAGGG	GGAGTACTGT	CGCAAGGCTG	[598]
TLV	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCAGGGG	GGAGTACTGT	CGCAAGGCTG	[600]
TPL	CGTCGATCAA	GAATGARASC	GTGGGGAGCA	RAAATGATCA	GACACCGTCG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCAGGGG	GGAGTACTGT	CGCAAGGCTG	[598]
LBU	CGTCGATCAA	GAATGTGAGT	GTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCCATG	GGTTCAGGG	GGAGTACTGT	TGCAAGGCTG	[587]
DAA	CGTCGATCAA	GAATGAGAGT	GTGGGGAGCA	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TCAACTGTGC	CGGCCGAGAA	ATCGCCATTG	GGTTCAGGGG	GAAGTACTGT	CGCAAGGCTG	[599]
LAM	CGTCGATCAA	GAATGTGAGT	GTGGGGAGCA	AAGATGATCA	GACACCGTTG	TAGTCCACTG	TAAACCATGC	CAGCCGGGAA	ACCGCCTATG	GGTTCAGGGG	GGAGTACTGT	CGCAAGGCTG	[600]
LTR	CGTCGATCAA	GAATGTGAGT	GTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TGAACCATGC	CAGCCGAGAA	ATCGCCTTCG	GGTTCAGGGG	GGAGTACTGT	CGCAAGGCTG	[564]
EUGLO	CGTCGATCAA	GAATGAGAGT	TCGGGGAGCA	AAGATGATCA	GACACCGTCG	TAGTCCGCTG	TAAACCATGC	CGGCCGAGAA	ATCGCCTGTG	GGTTCAGGGG	GGAGTACTGT	CGCAAGGCTG	[590]
EVI	CGTCGATCAA	GGATGAGAGT	CCGGGGAGCC	AAGATGATCA	GACACCGTCG	TAGTCCGCTG	TAAACCGTGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCAGGGG	GGAGTACATT	CGCAAGGCTG	[599]
LSGc38	CGTCGATCAA	GAATGTGAGT	GTGGGGAGCA	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TCAACTATGC	CGGCCGAGAA	ATCGCCTGTG	GGTTCAGGGG	GGAGTACTGT	CGCAAGGCTG	[564]
CPI	CATCGATCAA	GAATGAAAGC	GTGGGGAGCA	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCAGGGG	GAAGTACTGT	CGCAAGGCTG	[590]
CSK	CATCGATCAA	GAATGAAAGC	GTGGGGAGCA	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACCATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCAGGGG	GAAGTACTGT	CGCAAGGCTG	[590]
TET	CGTCGATCAA	GAACGAGAGT	GTGGGGAGCG	AAGATGATCA	GATACCGTCG	TAGTCCACCG	TAAACAATGT	CGGCCGAGAA	ATCGCTTGTG	GGTTCAGGGG	GGAGTACTGT	CGCAAGGCTG	[590]
EGN	CGTCGATCAA	GGATGAGAGT	CTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACGATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCAGGGG	GGAGTACCGT	CGCAAGGCTG	[600]
TVS	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCRCTG	TAAACCATGG	CAGCCGAGAA	ATCGCCTTTG	GGTTCAGGGG	GGAGTACTGT	AGCAAGGCTG	[600]
SNB	CGTCGATCAA	GAATGAGAGC	TCGGGGAGCG	AAGATGATCA	GACACCATCG	TAGTCCGCTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCCGGGG	GGAGTACGGT	CGCAAGGCTG	[600]
EHS	CGTCGATCAA	GAATGAGAGT	TCGGGGAGCA	AAGATGATCA	GACACCGTCG	TAGTCCGCTG	TAAACGATGC	CGGCCGAGAA	ATCGCCTGTG	GG-TCAGGGG	GGAGTACTGT	CGCAAGGCTG	[599]
GRC	CGTCGATCAA	GAATGAGAGT	TCGGGGAGCA	AAGATGATCA	GACACCGTCG	TAGTCCGCTG	TAAACGATGC	CGGCCGAGAA	ATCGCCTGTG	GGTTCAGGGG	GGAGTACTGT	CGCAAGGCTG	[590]
STB	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATTA	GACACCATCG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCCGGGG	GGAGTACGGT	CGCAAGGCTG	[600]
MUC	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGGT	CGCAAGGCTG	[599]
CVL	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGGT	CGCAAGGCTG	[564]
SOA	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATCA	GACACCATCG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCCGGGG	GGAGTACGGT	CGCAAGGCTG	[600]
SBR	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATCA	GACACCATCG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTGTG	GGTTCCGGGG	GGAGTACGGT	CGCAAGGCTG	[600]
CEL	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGGT	CGCAAGGCTG	[600]
CUA	CGTCGATCAA	GAATGAGAGC	GTGGGGAGCG	AAGATGATCA	GACACCGTCG	TAGTCCACTG	TAAACTATGC	CAGCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGGT	CGCAAGGCTG	[600]

	610	620	630	640	650	660	670	680	690	700	710	720	
HEXIN	AAACTTGAAG	GTATTGACGG	AAAGATACCA	CAAGACGTGG	AGTTGCGGCT	CAATTTGACT	CAACACGCAA	AT-CTTACCA	GACCCAGATG	-TACGGAGCA	ACAGATTAAG	AGTTCTTTTG	[693]
TRIPY	AAACTTAAAG	AAATTGACGG	AAGGGCACCA	CCAGGTGTGG	AGTTCGCGCT	TAATTTGACT	CAACACGGGG	AAACTTACCA	GGTCCAGACA	GGAAGGATTG	ACAGATTGAT	AGCTCTTTCT	[714]
MALJA	AAACTTAAAG	GAATTGACGG	AAGGGCA-CA	CCAGGAGTGG	AGTTCGCGCT	TAATTTGACT	CAACACGGGG	AAACTTACCA	GGTCCAGACA	GGAAGGATTG	ACAGATT-TC	AGCTCTTTCT	[709]
MALCA	AAACTTAAAG	GAATTGACGG	AAG-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[613]
DINEX	AAACTTAGAG	AAATTGACGG	AAGGGCACCA	CCAGATGTGG	AGTTCGCGCT	CAATTCGACT	CAACACGGGG	AAACTTATCA	GATTCCGACA	GTGAGGATTG	ACAGATTGAT	AGCTCTTTCT	[713]
MONSP	AAACTTAAAG	AAATTGACGG	AAGGGCACCA	CCAGGTGTGG	AGTTCGCGCT	TAATTTGACT	CAACACGGGG	AAACTTACCA	GGTCCAGACA	GGAAGGATTG	ACAGATTGAT	AGTTCTTTCT	[717]
STRST	AAACTTAAAG	AAATTGACGG	AAGGGCACCA	CCAGGTGTGG	AGTTCGCGCT	TAATTTGACT	CAACACGGGG	AAACTCACCA	GGTCCAGACA	GGAAGGATTG	ACAGATTGAT	AGCTCTTTCT	[714]
TRIVA	AAACTTGAAG	GAATTGACGG	AAGGGCA-CA	CCAGGGGTGG	AGTTCGCGCT	TAATTTGAAT	CAACACGGGG	AAACTTACCA	GGACCAGATG	TTTATGACTG	ACAGGCTTCG	GG-TCTTTCA	[689]
SPIVO	AAACTTGAAG	GTATTGACGG	AGGC-TACCA	CCAGACGTGG	AGTTGCGGCT	TAATTTGACT	CAACGCGCCA	AT-CTTACCTA	GACCTGATG	-TACTGTACG	TCAGATTGAT	AGTTCTTACA	[685]
ANDIN	AAACTTAAAG	GAATTGACGG	AAGGGCACCA	CCAGGAGTGG	AGTTCGCGCT	TAATTTGACT	CAACACGGGG	AAACTTACCA	GGTCCAGACA	GGAAGGATTG	ACAGATTGAC	AGCTCTTTCT	[700]
ANDGO	AAACTTAAAG	GAATTGACGG	AAGGGCACCA	CCAGGAGTGG	AGTTCGCGCT	TAATTTGACT	CAACACGGGG	AAACTTACCA	GGTCCAGACA	GGGAGGATTG	ACAGATTGAT	AGCTCTTTCT	[714]
SECEC	AAACTTAAAA	GAATTGACGG	AAGGGCACCA	CCAGGAGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGA	AAACTCACCA	GGTCCAGACA	GGGAGGATGG	ACAGATTGAG	AGCTCTTTCT	[716]
RECAM	AAACTTAAAG	GAAHTGACGG	AACCTTACCA	CCAATGGAGG	AGTTCGCGCT	TAATTCGACT	CAACTCGGGA	AACCTTACCA	GGTCCAGACA	GGGAGGATGG	ACAGATTGAG	AGCTCTTTCT	[715]
HISAR	AAACTTAAAG	GAATTGACGG	AAGAGCACGA	CAAGGCGAGG	AGTTCGCGCT	TAATTCGACT	CAACTCGGGG	AAACTTACCA	GAGCCAGACA	GGGAGGATGG	ACAGATTGAG	AGCTCTTTCT	[715]
JAKLI	AAACTTAAAG	GAATTGACGG	AAGGGCACCA	CCAGGAGTGG	AGTTCGCGCT	TAATTTGACT	CAACACGGGG	CAACTTACCA	GGTCCAGACG	GGGAGGATGG	ACAGATTGAT	AGCTCTTTCT	[716]
JAKBA	AAACTTAAAG	GAATTGACGG	AAGGGCACCA	CCAGGAGTGG	AGTTCGCGCT	CAATTTGACT	CAACACGGGG	CAACTTACCA	GGTCCAGACA	GGGAGGATGG	ACAGGCTGAC	AGCTCTTTCT	[597]
STALI	AAACTTAAAG	GAATTGACGG	AAAGGCACCA	CCAGGAGTGG	AGTTGCGGCT	TAATTCGACC	CAACACGGGG	AAACTCACGA	AGGCAAGATG	GATAGGATTG	ACAAAGTT-AA	AGACCTTTCT	[323]
NAEGR	AAACTTAAAG	GAATTGACGG	AAAGGCACCA	CAAGATGTGG	AGTTCGCGCT	TAATTTGACT	CAACACGGGG	AAACTCACCA	GGTCCAGACA	AGTTTGATTG	ACAGGTTAAT	AGCCCTTTCT	[710]
NAEAN	AAACTTAAAG	GAATTGACGG	AAAGGCACCA	CCAGGAGTGG	AGTTGCGGCT	TAATTCGACT	CAACACGGGG	AAACTCACCA	GGTCCAGACA	AGTTTGATTG	ACAGGTTAAT	AGCCCTTTCT	[711]
ACRRO	AAACTTAAAG	SAATTGACGG	AAAGGCACCA	CCAGGAGTGG	AGTTGCGGCT	TAATTCGACT	CAACGCGGGA	AAACTCACCA	GGTCAGGACA	GTAATGATTG	ACAGGTTGAT	AGTCTTTTCT	[709]
SAWMA	AAACTTAAAG	GAATTGACGG	AAATTCACCA	CCAGGAGTGG	AGTTCGCGCT	TAATTCGACT	CAACGCGTGA	AAACTTACCA	GATCCGAATT	GTAAGATTG	ACAGGTTAAT	AGCTTTTT-CT	[698]
BST	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACTTTACCA	GATCCGGACA	GTGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[716]
RBS	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACTTTACCA	GATCCGGACA	GTGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[703]
RMN	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGATGTGA	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACTTTACCA	GATCCGGAAT	GTAAGGATTG	ACAGATTGAT	TGTTCTTTCT	[716]
DMT	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAGTTTGACT	CAACACGGGG	AACTTTACCA	GATCCGGACA	GTAAGGATTG	ACAGATTGAT	TGTTCTTTCT	[718]
PERAM	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CCAGACGTGG	AGTTGCGGTT	TAATTTGACT	CAACACGGGA	AACCTTACCA	GGACAGGACA	GGAAATGATTG	ACAGATTGAT	AGCTCTTTCT	[650]
LEIMA	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[716]
LEIIN	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[700]
LEIBR	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[716]
LEITA	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[716]
LEPCO	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[716]
LEPBI	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[714]
LEPPO	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[716]
HERMU	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[715]
HERPE	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	ATGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[715]
BLACU	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	GGGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[717]
TRYBR	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	GTGAGGATTG	ACAGATGGAG	TGTTCTTTCT	[718]
TRYBE	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	GTGAGGATTG	ACAGATGGAG	TGTTCTTTCT	[718]
TRYAV	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	GTGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[716]
TRYCR	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGA	AGTTCGCGTT	TAATTTAAT	CAACACGGGG	AACCTTACCA	GATCCGGACA	GTGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[717]
TRYRA	AAACTTAAAG	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGTTCGCGTT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GATCCGGACA	GTGAGGATTG	ACAGATTGAG	TGTTCTTTCT	[717]
ENTO	AAACTTAAAG	GAATTGACGG	ACTCGCACCA	CAAGCAGTGG	ATTTGCGGCT	TAATTTGACT	CAACGCCCAA	AAACTTACTT	GGACAGAA	CTGGTGATG	ACGGATTGA-	TGATCCTTCA	[709]
PL0g	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGA	AACCTCGCCA	GGTCCGGA	CTGAGGATTG	ACAGATTGAG	AGCTCTTTCT	[715]
PERTR	AAACTTAAAG	GAATTGACGG	AGTGACACCA	CAAGGTGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GGTCAGGACG	GTTAGGATTG	ACAGACTGTG	AGCTCTTTCT	[717]
ESL	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGTGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AACCTTACCA	GGTCAGGACG	GCTAGGATTG	ACAGATTGAG	AGCTCTTTCT	[719]
EXS	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGTGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GTCTGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]
SAC	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACA	GCCAGGATGG	ACAGACTGCT	AGCTCTTGCT	[684]
ECL	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	ACCTGGATGG	ACAGATTGAG	AGCTCTTTCT	[718]
TSM	GAACCTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGA	AATGTTACCA	GGTCAGGACC	GCCAGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]
POR	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CCAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GTCGGGATGG	ACAGATTGAG	AGCTCTTTCT	[710]
TVI	GAACCTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCCGGGATGG	ACAGATTGAG	AGCTCTTTCT	[709]
MAA	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AACGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[718]
PSSc18	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACA	GCCGGGATGG	ACAGATTGAG	AGCTCTTTCT	[710]
EDS	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACA	GCCGGGATGG	ACAGACTGAG	AGCTCTTTCT	[718]
MOU	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AAGGTTACCA	GGTCAGGACA	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]
PYR	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AAGGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[719]
LVO	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCCAGGATGG	ACAGATTGAA	AGCTCTTTCT	[719]

	610	620	630	640	650	660	670	680	690	700	710	720	
TEA	AaAcTtAaAg	GaAtTgAcGg	AaTgGcAcCa	CaAgGcGtGg	AaTtGcGgCt	TaAtTtGaCt	CaAcAcGgGg	AaTgTtAcCa	GgTcAgGaCa	GtCaGgAtGg	AcAgAtTgAg	AgCtCtTtCt	[720]
EGR	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACA	GCCGGGATGG	ACAGATTGAG	AGCTCTTTCT	[719]
EPE	AAACTTAAAG	GAATTGACGG	ATTGGCACCA	CCAGGCGTGG	AGTTGCGGCT	TAATTTGATT	CAACACGGGA	AATGTTACCA	GGTCAGGACG	GCTGGGATTG	ACAGATTGAA	AGCCCTTTCT	[718]
TLV	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CMAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]
TPL	AAACTTAAAG	GAATTGACCG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[718]
LBU	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	NGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAA	AGCTCTTTCT	[707]
DAA	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGACTGAG	AGCTCTTTCT	[719]
LAM	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]
LTR	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[684]
EUGLO	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACGCGGGG	AATGTTACCA	GGTCAGGACG	ACTGGGATTG	ACAGATTGAG	AGCTCTTTCT	[710]
EVI	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[719]
LSGc38	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGTGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCC	[684]
CPI	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AACGTTACCA	GGTCAGGACG	GCTGGGATTG	ACAGATTGAG	AGCTCTTTCT	[710]
CSK	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AAAGTTACCA	GGTCAGGACG	GCTGGGATTG	ACAGATTGAG	AGCTCTTTCT	[710]
TET	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATTG	ACAGATTGAG	AGCTCTTTCT	[710]
EGN	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGTGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AACGTTACCA	GGTCAGGACA	GCCGGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]
TVS	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTTGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]
SNB	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACA	GCCAGGATGG	ACAGACTGCT	AGCTCTTGCT	[720]
EHS	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACGCGGGG	AATGTTACCA	GGTCAGGACG	ACTGGGATTG	ACAGATTGAG	AGCTCTTTCT	[719]
GRC	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACGCGGGG	AATGTTACCA	GGTCAGGACG	ACTGGGATTG	ACAGATTGAG	AGCTCTTTCT	[710]
STB	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACA	GTCAGGATGG	ACAGACTGCT	AGCTCTTGCT	[720]
MUC	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGACTGAG	AGCTCTTTCT	[719]
CVL	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[684]
SOA	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACA	GCCAGGATGG	ACAGACTGCT	AGCTCTTGCT	[720]
SBR	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACA	GCCAGGATGG	ACAGACTGCT	AGCTCTTGCT	[720]
CEL	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]
CUA	AAACTTAAAG	GAATTGACGG	AATGGCACCA	CAAGGCGTGG	AGTTGCGGCT	TAATTTGACT	CAACACGGGG	AATGTTACCA	GGTCAGGACG	GCTGGGATGG	ACAGATTGAG	AGCTCTTTCT	[720]

	730	740	750	760	770	780	790	800	810	820	830	840	
HEXIN	CGATTGGGCA	GTTGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGTGTAACCT	GTCCTTATT	GCCTTAACGA	GCGAGACAAC	CACCC-----	TGGGGGGAGC	GAAYGGGTTG	CTCTAGCAGG	[808]
TRIPY	TGATTCTGTG	GGTGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGTTAACGA	ACGAGACCTC	AGCCTCTTCT	TAGAGGGACG	GAAGTTGAGG	CAATAACAGG	[834]
MALJA	TGATTCTATG	GGTGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGATAACGA	ACGAGACCTT	AACCTCTTCT	TAGAGGGACG	GAAGATGAGG	CAATAACAGG	[829]
MALCA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[613]
DINEX	TGATTCCGTG	GGTGGTGGTG	CATGGTCGTT	CCTAGTTCGT	GGGTGGGCTT	GTCCTTAATT	GCGATAACGA	ACGAGACCTC	GGTCTTTCT	TAGAGAGACG	GAAG-TGAGG	CAATAACAGG	[832]
MONSP	TGATTCCGTG	GGTGGTGGTG	CATGGTCGTT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGATAACGA	ACGAGACCTC	AACCTCTCTC	TCGGGGGAGA	GTTGTTGAGA	TCTTTTAATG	[837]
STRST	TGATTCCATG	GGTGGTGGTG	CATGGTCGTT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGATAACGA	ACGAGACCTC	AACCTTTTCT	TAGAGGGACG	GAAGTTGAGG	CAATAACAGG	[834]
TRIVA	GGATATTGTT	TTTGGTGGTG	CATGGCCGCT	GGTGGTGCGT	GGGTTGACCT	GTCGTTGATT	CAGCTAACGA	GCGAGATTAT	CGCCAGTTCT	AATTGGGACG	GAAGAGGTAG	CAATAACAGG	[809]
SPIVO	TGAATGAGCA	GGTGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGTGTGAACCT	GTCCTTCAAT	CCGTTAACGA	GCGAGATGGA	CTCCG----	CTGCGGTATT	GAGAGGTCCG	CGCTAGCAGG	[800]
ANDIN	TGATTCTATG	GGTGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGATAACGA	ACGAGACCTT	CACC-CTTCT	TAGAGGGACG	GAAGTTGAGG	CAATAACAGG	[819]
ANDGO	TGATTCTATG	GGTGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGATTGATTT	GTCGTTAATT	CCGATAACGA	ACGAGACCTT	AACCTCTTCT	TAGAGGGACG	GAAGTTGAGG	CAATAACAGG	[834]
SECEC	AGATTCTATG	GGTGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGTTAACGA	ACGAGACCGA	AATCTCTTCT	TAGAGAGATT	GAACCTTCGG	CCAAAAACAGG	[836]
RECAM	TGATTCTATG	GGTGGTGGTG	CATGGCCGCT	TCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGATAACGA	ACGAGACCTC	TGCTCTTCT	TAGAGGGATT	GAACACAGAG	CTAGAACAGG	[835]
HISAR	TGATTCTATG	GATGGTGGTG	CATGGCCGCT	TTTAGTTCGT	GGAGTGATCT	GTCGTTAATT	CCGATAACGA	ACGAGACCTC	TGCTCTTCT	TAGAGGGATT	GAACACAGAG	CCAGAACAGG	[835]
JAKLI	TGATTCTGTG	GGTGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGTTAACGA	ACGAGACCGA	AATCTCTTCT	TAGAGAGATT	GAACCTTCGG	CCAAAAACAGG	[836]
JAKBA	AGATTCCATG	GATGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGAGTGATTT	GTCGTTGATT	CCGTTAACGA	ACGAGACCTC	CGTCTCTTCT	TAGAGAGACG	GAAGCGGAGG	CAAAAAACAGG	[717]
STALI	TGATTCTACA	AGTGGGGGTG	CATGGCCGCT	CCTAGTTCGT	GGGTTAACTT	GTCGTTAATT	CCGATAACGA	GCAAGACCTC	GGCT-CTTCT	TAGCAGGACG	GAAGCCGAGG	CAATAACAGG	[442]
NAEGR	TGATTGTGTG	GTGGGTAGTG	CATGGCCGCT	TCCAGTTCGT	GGAGTGATCT	GTCGTTAATT	CCGATAACGA	ACGAGACCTA	AGCTCTTCT	TAAAGGGACG	GAAGATTAGG	CCATAACAGG	[830]
NAEAN	TGATTGTGTG	GTGGGTAGTG	CATGGCCGCT	TCCAGTTCGT	GGAGTGATCT	GTCGTTAATT	CAGATAACGA	ACGAGACCTA	AGCTCTTCT	TAAAG-GACG	GAAGATTAGG	CCATAACAGG	[830]
ACRRO	CGATTCTTTG	GATGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CAGTTAACGA	ACGAGACCTT	AACCTCTTCT	TAGAGGAACG	GAAGTTAAGG	CAATAACAGG	[829]
SAWMA	TGATTCTATA	AAGAGTGGTG	CATGGCCGCT	CTCAGTTCGT	GGATTAATCT	GTGTTAACGA	ACGAGACCTT	ATCTACTTCT		TAGTAAACG	GA-A-TAAGG	CATATAACAGA	[816]
BST	CGATCCCTTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[836]
RBS	CGATCCCTTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	ATCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[823]
RMN	CGATTCTCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAATTTGGG	CAACAGCAGG	[836]
DMT	CGATCTCTTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[838]
PERAM	TGATTCTGTT	GACGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GAAGTGATTT	GTTGTTGATT	CCTTTAACGG	ACGAGACCCA	ATCGCTTCT	TCGCGGCATT	GGCATTGGG	CAACAGCAGG	[770]
LEIMA	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAATTTGGG	CAACAGCAGG	[836]
LEIIN	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAATTTGGG	CAACAGCAGG	[820]
LEIBR	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAATTTGGG	CAACAGCAGG	[836]
LEITA	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAATTTGGG	CAACAGCAGG	[836]
LEPCO	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAATTTGGG	CAACAGCAGG	[836]
LEPBI	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAATTTGGG	CAACAGCAGG	[834]
LEPPO	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAATTTGGG	CAACAGCAGG	[836]
HERMU	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[835]
HERPE	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[835]
BLACU	CGATTCCCTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[837]
TRYBR	CGATCCCTTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[838]
TRYBE	CGATCCCTTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[838]
TRYAV	CGATCCCTTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[836]
TRYCR	CGATCCCTTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	CTGCAGGATT	GAGATTTGGG	CAACAGCAGG	[837]
TRYRA	CGATCCCTTG	AATGGTGGTG	CATGGCCGCT	TTTGGTCCGT	GGAGTGATTT	GTTGTTGATT	CCGTCAACGG	ACGAGATCCA	AGCTGCTTCT	TTGCGGGATT	GAGATTTGGG	CAACAGCAGG	[837]
ENTO	TAAGTGTGTG	AAGGGTGGTG	CATGGCCACA	CGAAGTTCGT	GGGATGACCT	GTCCTTAATT	GCGATAACGG	GC-AGACCTT	CACCTGCTCT	AACCTTACAGG	TC-AAGGGGA	CTAATGTAGG	[827]
PLOg	TGATCGTGAG	GGTGGTGGTG	CATGGCCGCT	CTGGGTTCGT	GGAGTGATTT	GTCCTTAATT	GCGATAACGA	ACGAGATCAC	AACCTAGCCT	TAGAGGGACG	GAAGATGTGC	CTTTCACAGG	[835]
PERTR	TGATCCTGTG	GAGGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGATTTGGCT	GTCGTTAATT	CCGTAACGA	ACGAGAACTC	TACCTCTTCT	GCAAGGGACT	GAGGTAGAGT	CAACAGCAGG	[837]
ESL	TGATCCTGTG	GAGGGTGGTG	CATGGCCGCT	CCTAGTTCGT	GGAGTGATTT	GTCGTTAATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[839]
EXS	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TACCTTCTCT	CTGAGGCGCG	CATGCAGAGC	CAACAGCAGG	[840]
SAC	TGATTCCATG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[804]
ECL	TGATTCTGCG	GACGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[838]
TSM	TGATCCTTGG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
POR	TGATCCTTGG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[830]
TVI	TGATCCTTGG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[829]
MAA	TGATCCTTGG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TACCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[838]
PSSc18	TGATTCGTG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	TATGCAGAGC	CAACAGCAGG	[830]
EDS	TGATTCGGTG	GACGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[838]
MOU	TGATGCCTTG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TACCTCTTCT	CTGAGGTGCG	CATGCAGATA	CAACAGCAGG	[840]
PYR	TGATGCCTTG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TACCTCTTCT	CTGAGGTGCG	GATGCAGATA	CAACAGCAGG	[838]
LVO	TGATCCTTGG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[839]

	730	740	750	760	770	780	790	800	810	820	830	840	
TEA	TgAtTcAgTg	GaTgGtGgTg	CaTgGcCgCt	CcTgAtCgGt	GgAgTgAtTt	GtCgTtGaTt	CcGaTaAcGg	GtGaGaTaTc	TgCcTcTtCt	CtGaGgTgCg	CaTgCaGaGc	CaAcAgCaGg	[840]
EGR	TGATCCTGTG	GACGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[839]
EPE	TGATTCCTTG	GACGGTGGTG	CATGGCCGTT	CCTGATTGGT	GGAGTGATCT	GTCGTTGAGT	CCGATAACGA	GTGAGATCTC	TTTCTCTTCT	CCAAGGTGCG	TAAGCAGAGT	CAACAGCAGG	[838]
TLV	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
TPL	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[838]
LBU	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[827]
DAA	TGATCCTATG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[839]
LAM	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
LTR	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[804]
EUGLO	TGATCTTGTG	GACGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[830]
EVI	TGATCCTGTG	GACGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[839]
LSGc38	TGATTCCTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGTTAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[804]
CPI	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GAAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[830]
CSK	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GAAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[830]
TET	TGATCCTTTG	GACGGTGGTG	CATGGCCACT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGTAGG	[830]
EGN	TGATCCTGTG	GACGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
TVS	TGATTCCTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
SNB	TGATTCGATG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
EHS	TGATTCCTGTG	GACGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[839]
GRC	TGATTCCTGTG	GACGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[830]
STB	TGATCCAGTG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
MUC	TGATCCTTCG	GATGGTGGTG	CATGGCCGCC	CCGGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[839]
CVL	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[804]
SOA	TGATCCAGTG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
SBR	TGATCCAGTG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGATCTC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
CEL	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]
CUA	TGATCCTTCG	GATGGTGGTG	CATGGCCGCT	CCTGATTGGT	GGAGTGATTT	GTCGTTGATT	CCGATAACGA	GTGAGACATC	TGCCTCTTCT	CTGAGGTGCG	CATGCAGAGC	CAACAGCAGG	[840]

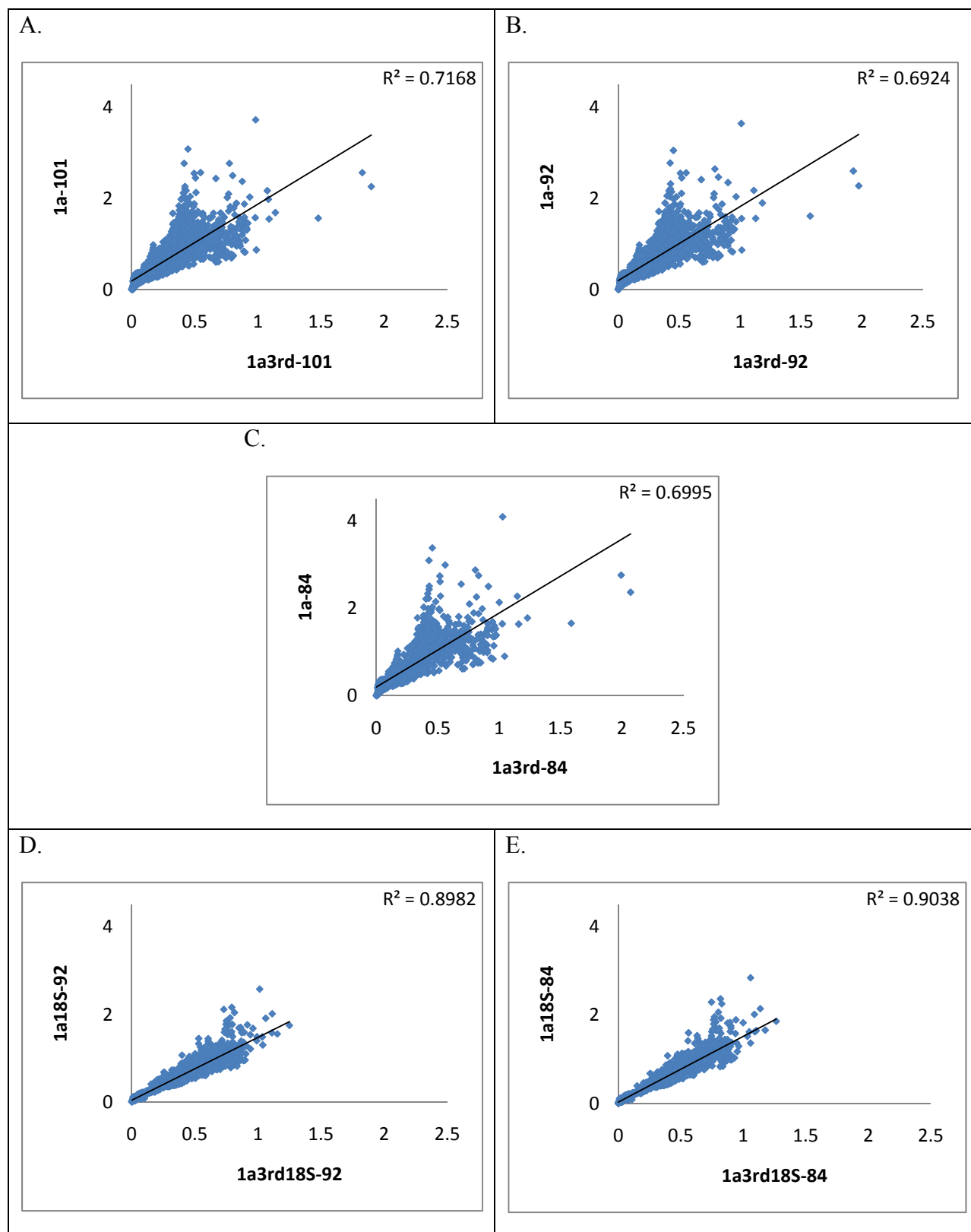
	850	860	870	880	890	900	910	920	930	940	950	960	
HEXIN	TCTGTGATGC	CTTAGACACT	CTGGCGCACG	CGTACTACAA	TGTCGGGCTC	ACAGTGGTAG	GGATCGGAGA	TTGGAATTTT	TCTAACGAGG	AATGTCTTGT	AGGTCTGCGT	TATTAGCGCC	[928]
TRIPY	TCTGTGATGC	CTTAGATGTT	CTGGCGCACG	CGCGCTACAC	TGATGTGCTC	AACATGATGG	GGATTGATCT	TTGCAATTAT	GGAAACGAGG	AATACCTAGT	AAGCGCAAGT	CAT-AGCTCC	[953]
MALJA	TCTGTGATGC	CTTAGATGTT	CTGGCGCACG	CGCGCTACAC	TGATGAGCTC	AACATGATGG	GGATAGACGT	TTRCAATTAT	CCGAACGAGG	AATTCTTAGT	AAGCGCGATT	CATCAGATCC	[949]
MALCA													[613]
DINEX	TCCGTATGTC	CTTAGACAAT	CTGACGACAG	CGTACTACTA	TAGCCGCGCTC	AACGTGCTGG	GGATCGGTTT	TTGTAATTTCT	GAAACGCGGG	AATATCTCGT	AAGTGCAGTT	CAT-AGACCC	[951]
MONSP	GTAAATATTA	AATATATATT	TTGAT-- -AG	CTCATTAAGC	TTTCAGGATG	GATTTTATTG	AAAT-TATCA	TTAAAATATT	TCGGATTGTA	ATTCTCTTTT	AGTTAATTAA	TTT-ATCAGT	[952]
STRST	TCTGTGATGC	CTTAGATGTC	CTGGCGCACG	CGCGCTACAA	TGATAGGGTC	AACATGCTGG	GGATTGATCT	TTGTAATTAT	GGAAACGAGG	AATACCTAGT	AAGTGCAGT	CATCAGCTCC	[954]
TRIVA	TCCGTGATGC	TTTAGATGCT	CTGGTGACAG	CGTGCTACAA	TGTTAGGATC	AATATAGTTG	GGATTGACGT	TTGTAATC-A	CGCAACGAGG	AATCCCTTGT	AAATGTGTGT	CAACAACGCC	[928]
SPIVO	TCTGTGATGC	CTTAGACACT	CTAGCGCACG	CGTACTACAC	TGTACGTGGC	AGCATGCTGG	GGATTGCGGG	TTGGAATTTCT	CCCAACGAGG	AATGTCTAGT	AATCATGAGT	CATTATCTCC	[920]
ANDIN	TCTGTGATGC	CTTAGATGTC	CTGGCGCACG	CGTGCTACAC	TGACGGATCC	AACATGATGG	GGATCGACCC	TTGTAAC--T	TGGAACGAGG	AATTCTTGT	AAGCGCAAGT	CAT-ATCTCC	[936]
ANDGO	TCTGTGATGC	CTTAGATGTT	CTGGCGCACG	CGCGCTACAC	TGACGGATGC	AGCATGATGG	GG-AAAGCCC	TTGCAATTAT	TGGAACGAGG	AATTCTTGT	AAGCGCAAGT	CATCAACTCC	[953]
SECEC	TCTGTGATGC	CTTCGATGTT	CTGGCGCACG	CGCGCTACAC	TGATGGACGC	AACGTGCTGG	GGATCGACCC	TTGCAATTTG	GGGAACGAGG	AATTCTTAGT	AGGCGCGGGT	CATCAGCCCC	[955]
RECAM	TCTGTGATGC	CTTCGATGTT	CTGGCGCACG	CGCGCTACAC	TGACGGAGGC	AACATGCTGG	GGATAGACCA	TTGCAATTAT	TGGAACGAGG	ATTCTGTTAGT	AATCGCGTGT	CATCATCGCC	[955]
HISAR	TCTGTGATGC	CTTCGATGTT	CTGGTGACAG	CGCGCTACAC	TGACGAAACC	AACATGCTGG	GGATAGACTA	TTGCAATTAT	TAGAACGAGG	AATGGCTAGT	AAGCGCGTGT	CAT-ATCGCC	[954]
JAKLI	TCTGTGATGC	CTTCGATGTT	CTGGCGCACG	CGCGCTACAC	CGGTGAGCTC	AACGTGATTT	GGATGACTCT	TTGCAATTTT	TGGAACGAGG	AATTCTTAGT	AAGCGCGAGT	CATCAGCTCC	[956]
JAKBA	TCTGTGATGC	CTTCGATGTC	CTGGCGCACG	CGCGCTACAC	TGATGAACGC	AACGTGTTTG	GGATAGACCA	TTGCAACTAT	TGGAACGAGG	AATTCTTAGT	AAGCGCGAGT	CAT-AACTCC	[836]
STALI	TCTGTGATGC	CTTAGATGCC	TTGAGACAGC	CGTACTACAA	TGTTAAAATA	AGCATGACAG	GGATTGGTCA	TTGGA---G	TGGAACGAGG	AATTCTCTGT	AATCGCAAGT	CATCATCTCC	[559]
NAEGR	TCTGTGATGC	CTTAGATGTC	CTGGTGACAG	CGTACTACAA	TAACGGTACC	AGCGTGACAG	GGATCGAGGA	TTGGAAC--A	TCCAACGAGG	AATTCTTAGT	AAGCGTGGTT	CATCATACCC	[948]
NAEAN	TCTGTGATGT	CTTAGATGTC	CTGGTGACAG	CGTACTACAC	TGTTAATACT	AACATGACAG	GGATAGCTGT	TTGGA---A	TCCAACGAGG	AATTCTTAGT	AATCGCGATT	CATCATATCC	[946]
ACRRO	TCTGTAATGC	CCTAGATGTC	CTGGAGCACG	CGTACTACAC	TGTTAATGTT	GAATTAATGAG	GAATTAATGAG	TTGTA---A	TCCAACGAGG	AATTCTTAGT	AATAGTAAAT	CATATTCTCC	[933]
SAWMA	TCTGTGATGC	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[956]
BST	TCTGTGATGT	CCTCAATGTT	CTGGTGACAG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGCCTCGT	AGGCGCAGCT	CATCAAACCC	[943]
RBS	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGCCTCGT	AGGCGCAGCT	CATCAAACCC	[956]
RMN	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGCCTCGT	AGGCGCAGCT	CATCAAACCC	[958]
DMT	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGCCTCGT	AGGCGCAGCT	CATCAAACCC	[888]
PERAM	TCCGTGATGT	CCTAAATGTC	CTGGGACACG	CGCACTACAA	TGACAAAAGTG	AATGTGCTTG	GGATCGAAGC	TTGCAATTAT	TCTAACGAGG	AATGTCTAGT	AGACGTGCTT	CATCAAAG--	[956]
LEIMA	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[940]
LEIIN	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[956]
LEIBR	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[956]
LEITA	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[956]
LEPCO	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[954]
LEPBI	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[956]
LEPPO	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[955]
HERMU	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACAGTTTTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[955]
HERPE	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACAGTTTTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[957]
BLACU	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGGA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[958]
TRYBR	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[958]
TRYBE	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[956]
TRYAV	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[957]
TRYCR	TCTGTGATGT	CCCCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GAACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[957]
TRYRA	TCTGTGATGT	CCTCAATGTT	CTGGGACACG	CGCACTACAA	TGTCAGTGAG	AACACAATTG	GGACCGAGTA	TTGCAATTAT	TGGAACGAGG	AATGTCTCGT	AGGCGCAGCT	CATCAAACCC	[957]
ENTO	TCTGTGATGC	CCCAGATGCC	CAAGCACACG	CGCGATACAA	TGGAGTTGGC	AGCATGGAAG	GGATTGCCGG	CAGTACA-CG	CCGAACGAGG	AATTGCTTGT	AAGTACCTGT	GATCAGCAC	[946]
PLOg	TCTGTGATGC	CGC-AATGTT	CTGGTGACAG	CGCACTACAC	TGGCATAGTC	AACGGACCGG	GGATAGGTGA	TTGCAATTAT	TCGAACGAGG	AATGCCAGT	AAGGGCGAGT	CATCAGTCC	[954]
PERTR	TCTGTGATGT	CCTAAATGTC	CTGGTGACAG	CGCACTACAC	TGGCAAGGAC	ACTTTGTGAG	GGATAGATGT	TTGCAATTAT	CCAAACGAGG	AATACCTAGT	AAGCATAGGT	CATCGGCCCC	[957]
ESL	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAC	TGATCGATGC	AAGGTGCTGG	GGATAGATGT	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AGGCGTTGGT	CACGAGCCCC	[959]
EXS	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACTCAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATACCTAGT	ATGCGTGGGT	CATCAGCCCC	[960]
SAC	TCTGTGATGT	CCTAGATGTT	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AATGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGTGAGT	CAT-AGCTCC	[923]
ECL	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAC	TGGCCCGGTG	AAGGACCCSS	SGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	CTCCCTCTGA	[957]
TSM	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGATTCTAGTG	AAGGCGATGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCGTGT	CAT-AGCACC	[959]
POR	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGGTCTGGTG	AATGTACTGG	GGATAGATGC	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AGACGTGGGT	CATCAGCCCC	[950]
TVI	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGATTCTAGTG	AAGGCGCTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCGTGT	CAT-AGCACC	[948]
MAA	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGATACAGTG	AAGGTCTCTG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AGGCGTGCGT	CATCAGCACC	[958]
PSSc18	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGGTCTGGTG	AATGTACTGG	GGATAGCTGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCTTGGGT	CAT-AGCCCC	[949]
EDS	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGATTCTAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AGGCGTGAGT	CATCAGCTCC	[958]
MOU	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGCC	AGGTCCTG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTTGT	AAGCGTGTTG	CAT-AGCGCC	[959]
PYR	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGCC	AGGTCCTG	GGATAGATGC	TTGCAACTGT	CTGAACGAGG	AATGCCTTGT	AAGGGTGTTG	CAT-AGTGCC	[957]
LVO	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACCCAGTG	AATATACTTG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AGGCGTGGGT	CAA-AGCCCC	[958]

	850	860	870	880	890	900	910	920	930	940	950	960	
TEA	TcTgTgAtGt	CcCaGaTgTc	CtGgCgCaCg	CgCaTtAcAt	TgAcAcAgTg	AaGgTaCtGg	GgAtAgAtGa	TtGcAaCtGt	CtGaAcGaGg	AaTgCcTaGt	AtGcTtGtGt	CaTcAgCaCc	[960]
EGR	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGGTGCAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGTGG	AATGCCTAGT	ATGCCTTGGT	CATCAGCCCC	[959]
EPE	TCTGTGATGT	CCTAGATGTC	CTGGCGCACG	CGCACTACAA	TGATACAATG	AAAGTGTCGG	GGATCGATGC	TTGCAACTGT	TCGAACGAGG	AATGCCTAGT	ATACGCACGT	CAT - ACCTCC	[957]
TLV	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AGGCGTGGGT	CATCAGCCCC	[960]
TPL	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AGGCGTGGGT	CATCAGCCCC	[958]
LBU	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGGCCCAATG	AATATACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AAGTGTGGGT	CATCAGCCCC	[947]
DAA	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AATGTACTGG	GGATAGAGCT	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCAGGT	CAT - AGCCCC	[958]
LAM	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGGTCCAGTG	AATGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AAGCGTGGGT	CAT - AGCCCC	[959]
LTR	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGATCTGGTG	AATGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCAGGT	CAT - AGCCCC	[923]
EUGLO	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGTCACAGTG	AGGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGTGG	AATGCCTAGT	ATGCCCTGGT	CATCAGCCCC	[950]
EVI	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACTCAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGTGG	AATGCCTAGT	AGGCGTTGGT	CATCAGCCCC	[959]
LSGc38	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGGCCCAAGT	AATGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATACCTAGT	AAGCGCAGGT	CAT - AGCCCC	[923]
CPI	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AAGGTACTTG	GGATAGAGCC	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCAGGT	CAT - AGCCCC	[949]
CSK	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AAAGTACTTG	GGATAGAGCC	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCCCAAGT	CAT - AGCCCC	[949]
TET	TCTGTGATGT	CCTAGATGTC	CTGGCGCACG	CGCACTACAC	TGACACAGTG	AACGTGCTGG	GGATAGATGA	TTGCAACTAT	TTGAACGAGG	AATGCCTAGT	AAGCGCAAGT	CATCAGCTCC	[950]
EGN	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AATGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGCGG	AATACCTAGT	ATGCGCAGGT	CATCAGCCCC	[960]
TVS	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	AGGCGTGGGT	CATCAGCCCC	[960]
SNB	TCTGTGATGT	CCTAGATGTT	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AATGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGTGAAGT	CAT - AGCTCC	[959]
EHS	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGTCACAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGTGG	AATGCCTAGT	ATGCCCTGGT	CATCAGCCCC	[959]
GRC	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGTCACAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGTGG	AATGCCTAGT	ATGCCCTGGT	CATCAGCCCC	[950]
STB	TCTGTGATGT	CCTAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGTGAAGT	CAT - AGCTCC	[959]
MUC	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGGG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCAAGT	CATCAGCTCC	[959]
CVL	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGGG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCAAGT	CAT - AGCTCC	[923]
SOA	TCTGTGATGT	CCTAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGTGAAGT	CAT - AGCTCC	[959]
SBR	TCTGTGATGT	CCTAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGTG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGTGAAGT	CAT - AGCTCC	[959]
CEL	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGGG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCAAGT	CATCAGCTCC	[960]
CUA	TCTGTGATGT	CCCAGATGTC	CTGGCGCACG	CGCACTACAT	TGACACAGGG	AAGGTACTGG	GGATAGATGA	TTGCAACTGT	CTGAACGAGG	AATGCCTAGT	ATGCGCAAGT	CATCAGCTCC	[960]

	970	980	990	1000	1010	1020	1030	1040	
HEXIN	TGTCTTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGACGGAT	GTCTGTGAGT	TGTCACGACG	AACCTGGAGG	AAGGAGAAGT	CGTAACAA [1016]
TRIPY	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGAAT	GATCATGAAA	TGTTTGGATA	AATCTAGAGG	AAGGAGAAGT	CGTAACAA [1041]
MALJA	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAT	GGTCATGAAG	TCTCCGGATA	AATCTAGAGG	AAGGAGAAGT	CGTAACAA [1037]
MALCA	-----	-----	-----	-----	-----	-----	-----	-----	[613]
DINEX	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGA-GAAT	GATCGTGAAA	TGCTCGGATA	AATCTAGAGG	AAGGAGAAGT	CGTAACAA [1038]
MONSP	T----GTGGT	GGAAGGCTAA	TGGGGCTCTG	GATCTAATTC	GAACACTTAG	AG--GGACTG	ATGCTTCTGG	GCCGCACGCG	CGCTACAC [1034]
STRST	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATAAAT	GATCATGAAA	TGTTCCGGATG	AATCTAGAGG	AAGGAGAAGT	CGTAACAA [1042]
TRIVA	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAT	GACTGTGAAA	TCACCGGATA	AATCTAGA--	-----	[996]
SPIVO	TGGCTCTTGT	ACACACCGCC	CGTCGCTTCT	ACCGATGTGT	GCT-TTGAGT	TCTATGGACG	AAGCTAGAGG	AAGGAGAAGT	CGTAACAA [1007]
ANDIN	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	A-CGA-GGAT	GGTCGTGAAA	TCTTCGGA--	-----	-----	[992]
ANDGO	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAT	GGTCGTGAAC	TCTTCGGATA	AACCTAGAGG	AAGGAGAAGT	CGTAACAA [1041]
SECEC	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAT	GATGGTGAAA	TCTCAGGATA	AACCTAGAGG	AAGGAGAAGT	CGTAACAA [1043]
RECAM	TGGAGTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAT	GATGGTGAAA	TCTTCGGATG	AATCTAGAGG	AAGGAGAAGT	CGTAACAA [1043]
HISAR	TGCTCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAT	GATGGTGAAA	TCTCCGGATA	AATCTAGAGG	AAGGAGAAGT	CGTAACAA [1042]
JAKLI	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAT	GATGGTGAAAG	ATTCTGGATA	AATCTAGAGG	AAGGAGAAGT	CGTAACAA [1044]
JAKBA	TGCCCTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAT	GATGGTGAAA	CCTCCGGATA	AATCTAGAGG	AAGGAGAAAC	C----- [917]
STALI	TG-----	-----	-----	-----	-----	-----	-----	-----	[561]
NAEGR	TGCCTTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAC	GAAGATGAAC	CTGGCGGATA	AATCTGTAGG	AAGGAAAAGT	CGTAACAA [1036]
NAEAN	TGCCTTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGGAC	GAAGATGAAC	CTGGCGGATA	AATCTGTAGG	AAGGAAAAGT	CGTAACAA [1036]
ACRRO	TGCCTTTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGAAT	GAGCGTGAAT	TTGACGGATG	AACCTGTAGG	AAGGAGAAGT	CGTAACAA [1034]
SAWMA	TGAAATTTGT	ACACACCGCC	CGTCGCTCCT	ACCGATGAAT	AAAAATAAAA	TTGATGGATA	AATCTGTAGG	AAGGAGAAGT	CGTAACAA [1021]
BST	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGTAACAA [1044]
RBS	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGA	TGATCGGACG	ATATTAGAGG	AAGCAAA--	----- [1020]
RMN	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTGCAGG	TGATCGGACG	ACATTAGAGG	AAGCAAAAAGT	CGTAACAA [1044]
DMT	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GCCTGTCAGG	TGATCGGACG	ACATTAGA--	-----	[1026]
PERAM	-----	-----	-----	-----	-----	-----	-----	-----	[888]
LEIMA	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAA----	----- [1032]
LEIIN	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	AT-----	-----	[1002]
LEIBR	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTA----	-----	[1022]
LEITA	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGTAACAA [1044]
LEPCO	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAA----	----- [1032]
LEPBI	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	C----- [1035]
LEPPO	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	C----- [1037]
HERMU	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	----- [1035]
HERPE	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAA----	----- [1031]
BLACU	TGCCATTTGT	ACACACCGCC	CGTCATTGTT	TCCGATGATG	GTGGTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGT----- [1040]
TRYBR	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGTAACAA [1046]
TRYBE	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGTAACAA [1046]
TRYAV	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	CGAAACAA [1044]
TRYCR	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAG-	----- [1036]
TRYRA	TGCCATTTGT	ACACACCGCC	CGTCGTTGTT	TCCGATGATG	GTGCTACAGG	TGATCGGACG	ATATTAGAGG	AAGCAAAAAGT	C----- [1038]
ENTO	TGCGAGTTGT	ACACACCGCC	CGTCGTTGCT	ATCGATTGCA	GCCGGAAAAGC	TTTACCTGTG	ATTCTAGAGG	AAGCAAAAAGT	CGTAACAA [1034]
PL0g	TGGCATTTGT	ACACACCGCC	CGTCGTTGTC	GTCGATGGCT	GTAGTAGAGG	CGTTGAGACA	ATATTAGAGA	CGATAAAAAGT	CGTAACAA [1042]
PERTR	TGCCACTTGT	ACACACCGCC	CGTCGTTGCT	ACTGATGGTA	TGGTCAGAGG	CTGGAAGACA	ATGTTAGAAG	AAGCAAAAAGT	CGTAACAA [1045]
ESL	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGATC	GCTGTAGAAC	CATCAGGATG	ATGCTAGAGG	AAGCAAAAAGT	CGTAACAA [1047]
EXS	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCA	TCTGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAAGT	CGTAACAA [1048]
SAC	TGCC-----	-----	-----	-----	-----	-----	-----	-----	[927]
ECL	AGAGGGTGGG	AAATCCTGCA	AGTCCGTTTG	TGAGAGCCGT	GCTGTATAGT	ATTGCAACTT	GTGTTTGTAC	CACCGCCCGT	CGTTGCTA [1045]
TSM	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGACA	GCTGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAAGT	CGTAACAA [1047]
POR	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTA	CGCGTAGAGC	CATCCGGACA	ATGTTAGAGG	AAGCAAAA--	----- [1028]
TVI	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGACA	GCTGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAA--	----- [1026]
MAA	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGAAA	GCAGTAGAGC	CATCAGGACA	ATGCTAGAGG	AAGCAAAAAGT	CGTAACAA [1046]
PSSc18	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCA	GCAGTAGAGC	CATCAGGACG	ATGTTAGAGG	AAGCAAAA--	----- [1027]
EDS	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCG	GCCGTAGAGC	CCTCCGGATG	ATGTTAGAGG	AAGCAAAA--	----- [1036]
MOU	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	ATGGCAGAGC	CATCCGGACA	ATGTTAGAGG	AAGCAAAAAGT	CGTAACAA [1047]
PYR	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	ATGGCAGAGC	CATCCGGACA	ATGTTAGAGG	AAGCAAAAAGT	CGTAACAA [1045]
LVO	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCA	GCCTAGAGC	CATCAGGACG	ATGTTAGAGG	AAGCAAAAAGT	CGTAACAA [1046]

	970	980	990	1000	1010	1020	1030	1040		
TEA	TgCcAtTtGt	AcACaCcGcC	CgTcGtTgCt	AcCgAtGaGa	GcAaCaGaGt	CaTcAgGaCa	AtGcTaGaGg	AaGcAaAaGt	CgTaAcAa	[1048]
EGR	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGACA	GCTGTAGAAC	CATCAGGATG	ATGCTAGAGG	AAGCAAAA--	-----	[1037]
EPE	TGCCAATTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTT	GTTGCAGAGA	CTTCAGGACT	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1045]
TLV	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGATC	GCTGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1048]
TPL	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGATC	GCTGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1046]
LBU	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTA	GCTGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1035]
DAA	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCA	GCTGTAGAGC	CATCAGGACA	ATGTTANNGG	AAGCAAAA--	-----	[1036]
LAM	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	GCAGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1047]
LTR	TGCC-----	-----	-----	-----	-----	-----	-----	-----	-----	[927]
EUGLO	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	GCTGTAGAGT	CATCAGGACA	ATGTTAGAGG	AAGCAAAA--	-----	[1028]
EVI	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGATA	GCTGTAGAAC	CATCAGGATA	ATGTTAG-AG	AAGCAAAAGT	CGTAACAA	[1046]
LSGc38	TGCC-----	-----	-----	-----	-----	-----	-----	-----	-----	[927]
CPI	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCA	GCAGTAGAGT	CATCAGGACA	ATGTTAGAGG	AAGCAAAA--	-----	[1027]
CSK	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCA	GCAGTAGAGT	CATCAGGACA	ATGTTAGAGG	AAGCAAAA--	-----	[1027]
TET	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGATA	GAAGTAGAGG	CATCAGGACA	ATGTTAGAGG	AAGCAAAA--	-----	[1028]
EGN	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTA	GCTGTAGAAC	CATCAGGATA	ATGTTAG-AG	AAGCAAAAGT	CGTAACAA	[1047]
TVS	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGACA	GCTGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1048]
SNB	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGACGGCA	GCTGTAGAGC	CATCAGGACA	ATGTTGAAGG	AAGCAAAAGT	CGTAACAA	[1047]
EHS	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	GCTGTAGAGT	CATCAGGACA	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1047]
GRC	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	GCTGTAGAGT	CATCA-----	-----	-----	-----	[1005]
STB	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTA	GCTGTAGAGC	CATCAGGACA	ATGTTAAAGG	AAGCAAAAGT	CGTAACAA	[1047]
MUC	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	GCAGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1047]
CVL	TGCC-----	-----	-----	-----	-----	-----	-----	-----	-----	[927]
SOA	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCA	GCTGTAGAGC	CATCAGGACA	ATGTTAAAGG	AAGCAAAAGT	CGTAACAA	[1047]
SBR	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGCA	GCTGTAGAGC	CATCAGGACA	ATGTTAAAGG	AAGCAAAAGT	CGTAACAA	[1047]
CEL	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	GCAGTAGAGC	CATCAGGACA	ATGTTAGAGG	AAGCAAAAGT	CGTAACAA	[1048]
CUA	TGCCATTTGT	ACACACCGCC	CGTCGTTGCT	ACCGATGGTG	GCAGTAGAGC	CATCAGGACA	ATGTTAG-AG	AAGCAAAAGT	CGTAACAA	[1047]

Figure 2.3. Saturation plots of EF1 α -included nucleotide alignments. Alignments with all codon positions included (y-axis) vs. their respective alignment with the third codon positions removed (x-axis). Pairwise ML distances were calculated within each data set individually in PAUP* based on empirical values directly from the data or from empirical estimations of invariant sites and gamma shape parameters as reported by RAxML. **A.** 1a-101 (1227bp) vs. 1a3rd-101 (818bp), **B.** 1a-92 (1230bp) vs. 1a3rd-92 (820bp), **C.** 1a-84 (1227bp) vs. 1a3rd-84 (818bp), **D.** 1a18S-92 (2238bp) vs. 1a3rd18S-92 (1828bp), **E.** 1a18S-84 (2275bp) vs. 1a3rd18S-84 (1866bp). A-C represent the single gene analyses of EF1 α with different OTU sampling and D-E represent the concatenation of EF1 α with 18S with different OTU sampling.



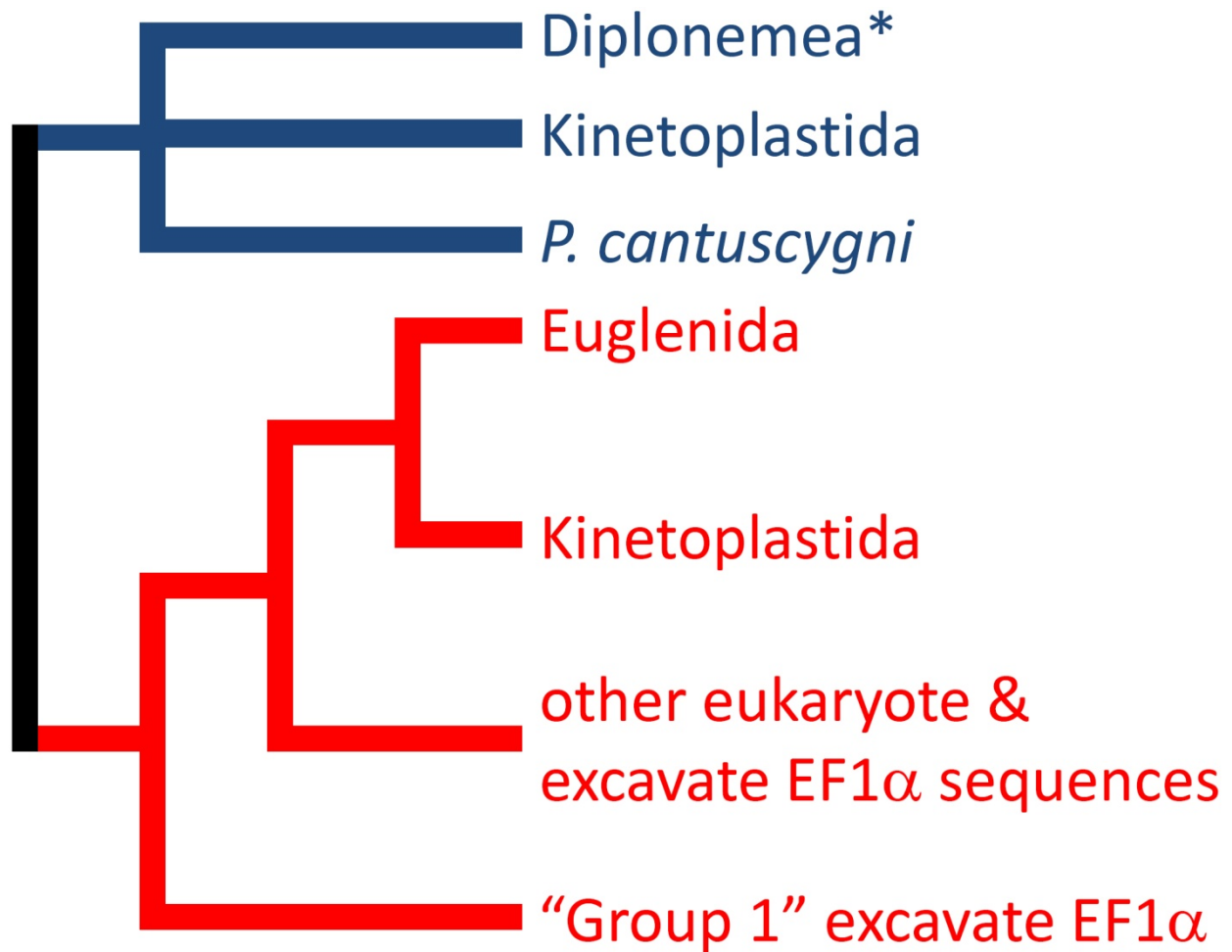


Figure 2.4. Cartoon distribution of excavate and euglenozoan organisms in the EF1 α and EFL gene tree based on Fig. 1.4 in Chapter 1 of this dissertation. EF1 α is the red portion of the cartoon and EFL is the blue portion. The asterisk represents the euglenozoan superclass that is only represented by EFL in this gene tree. Kinetoplastids are represented by both genes; however, they are recovered as monophyletic clades in both the EF1 α and EFL portions of the gene tree.

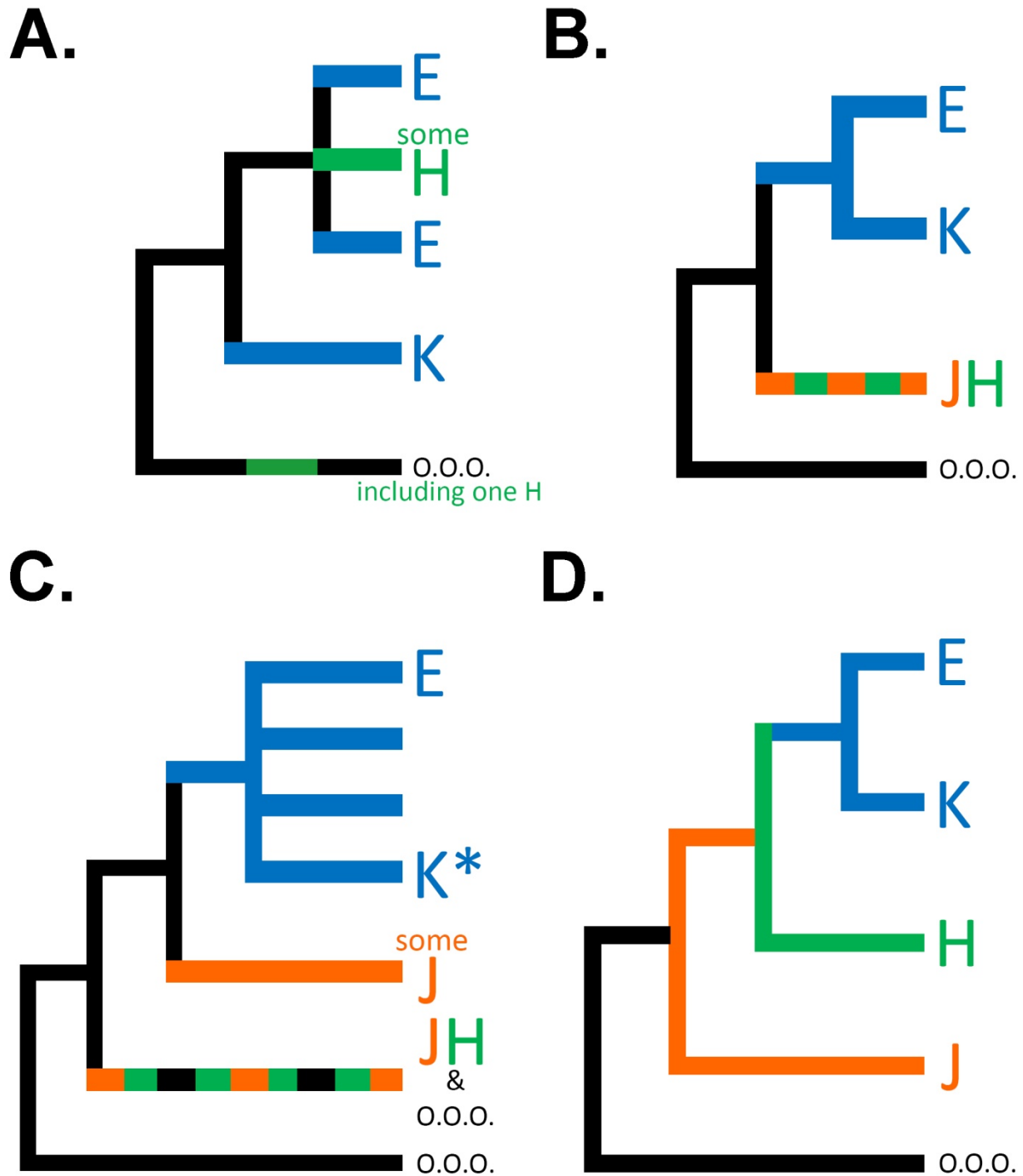


Figure 2.5. Cartoon summaries of single and concatenated gene ML and MLE analyses with 84 OTUs. Blue=Euglenozoa (E=Euglenida, K=Kinetoplastida and unlabeled euglenozoan branches=euglenid OTUs), green and H=Heterolobosea, o.o.o.=other excavate outgroup OTUs, orange and J=Jakobida,. **A.** Topology of 1a-84 analyses, **B.** Topology of 1a3rd-84 analyses, **C.** Topologies of 18S-84 analyses, **D.** MLE analysis of 1a3rd18S-84. ML and MLE analyses of 1a18S-84 and ML analysis of 1a3rd18S-84 was the same as this topology, except a long-branching euglenid OTU, *Entosiphon* sp., branches as basal to the Kinetoplastida.

Figure 2.6. RAxML ML tree from the 1a-101 alignment analysis. The left or top value is the RAxML bootstrap score and the MrBayes posterior probability is below or to the right. Only bootstraps $\geq 50\%$ and posterior probabilities ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bootstrap support from RAxML and 1.00 posterior probability from the MrBayes resulting tree topologies for the same alignment. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Clade names are to the right and if the clade was not recovered as monophyletic and the two places on the tree where the parts of the clade were found were not adjacent, then the clade names were numbered. Branches with hash marks are shown at one half their actual lengths. Euglenozoan OTUs are highlighted with blue. Checks for sequence placement in the tree vs. species tree and identical sequences was based on this alignment and stars denote sequences that were not used in further analyses.

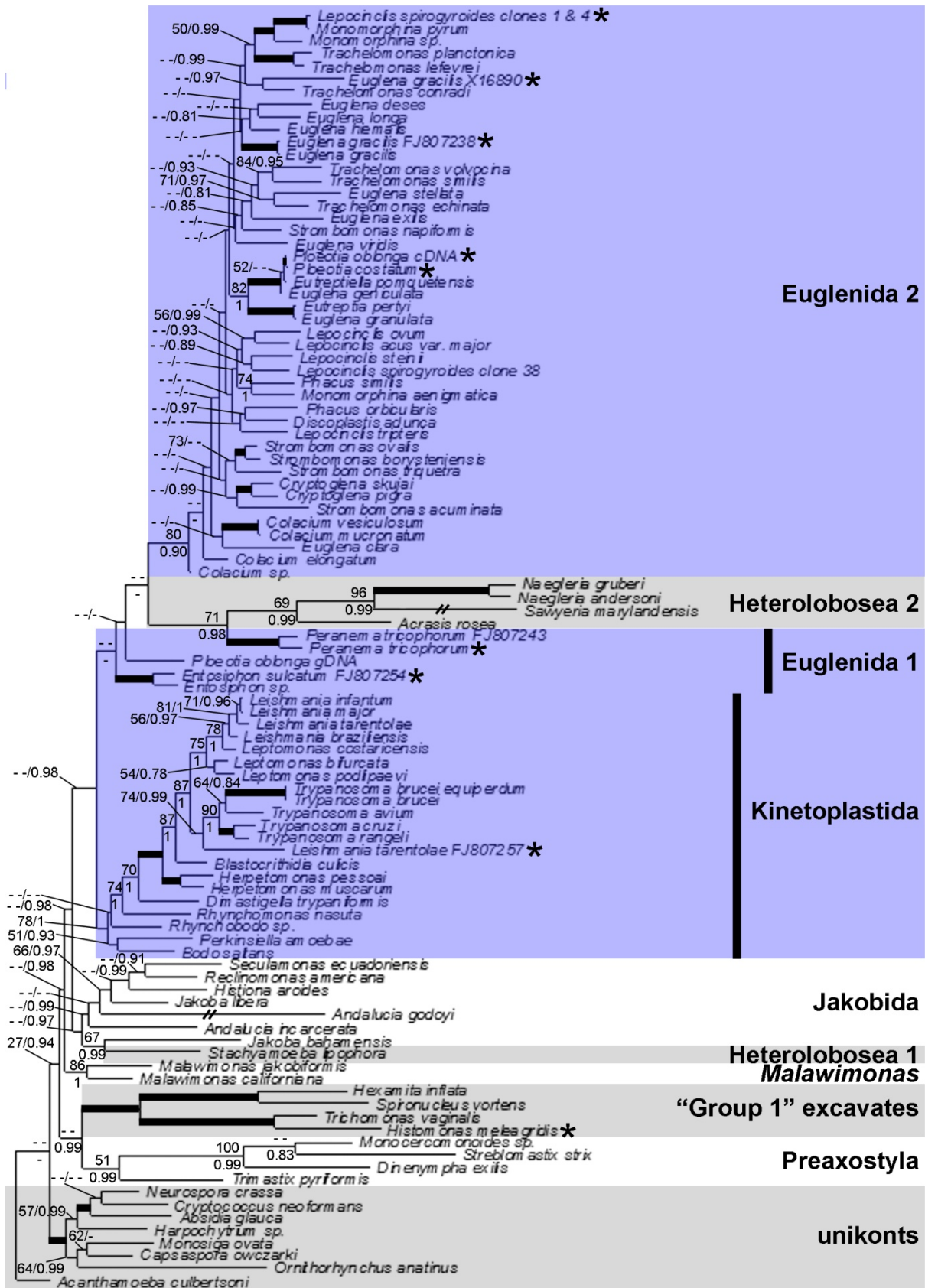


Figure 2.7. RAxML ML tree from the 1a3rd-92 alignment analysis. The left or top value is the RAxML bootstrap score and the MrBayes posterior probability is below or to the right. Only bootstraps $\geq 50\%$ and posterior probabilities ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bootstrap support from RAxML and 1.00 posterior probability from the MrBayes resulting tree topologies for the same alignment. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Clade names are to the right and if the clade was not recovered as monophyletic and the two places on the tree where the parts of the clade were found were not adjacent, then the clade names were numbered. The unikonts are not recovered as a clade in this analysis and was therefore broken into its supergroup members; Amoebozoa and Opisthokonta. The branch with hash marks is shown at one half its actual length. Euglenozoan OTUs are highlighted with blue.

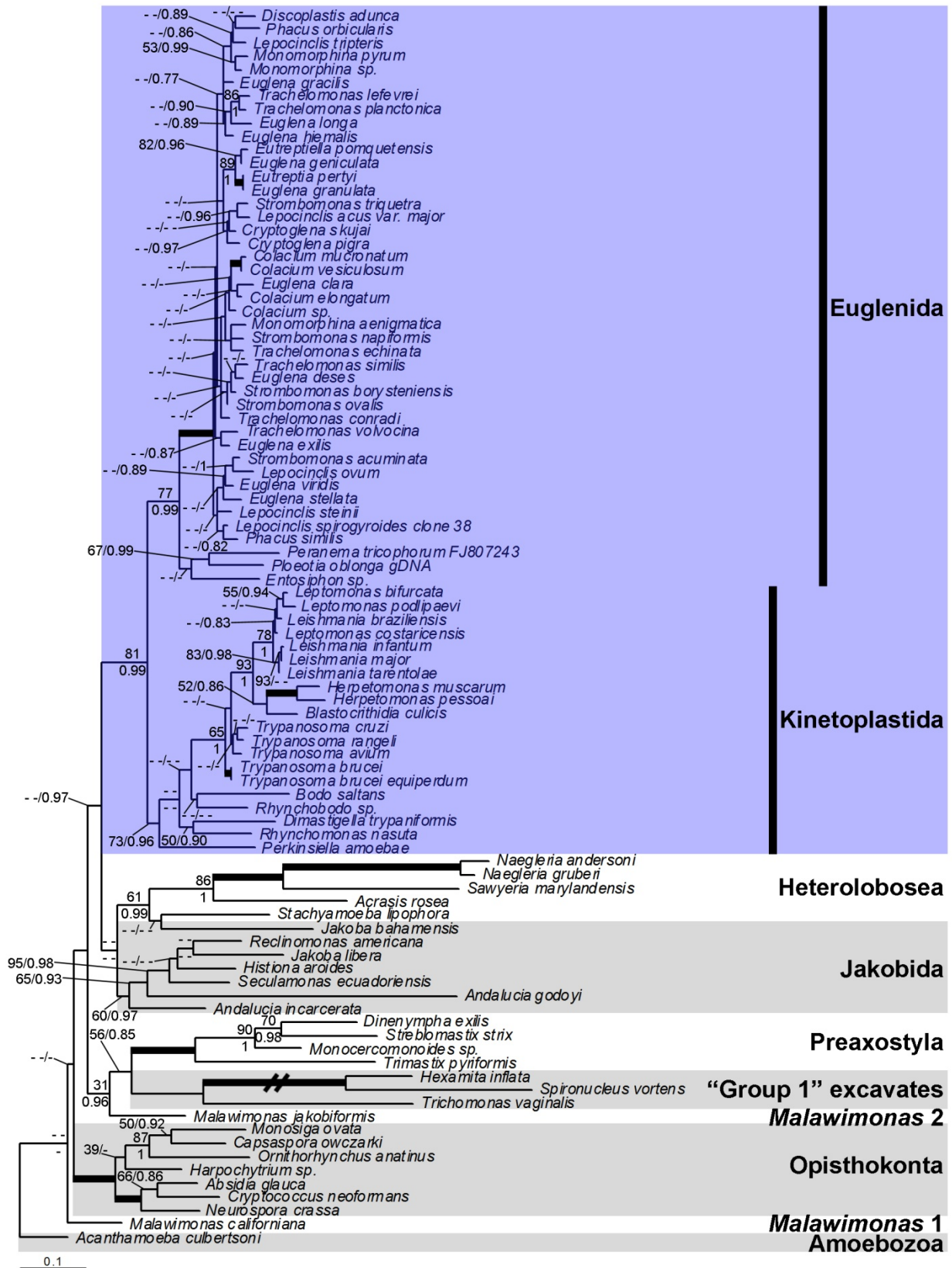


Figure 2.8. MrBayes consensus tree from the 18S-92 alignment analysis. The left or top value is the RAxML bootstrap score and the MrBayes posterior probability is below or to the right. Only bootstraps $\geq 50\%$ and posterior probabilities ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bootstrap support from RAxML and 1.00 posterior probability from the MrBayes resulting tree topologies for the same alignment. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Clade names are to the right and if the clade was not recovered as monophyletic and the two places on the tree where the parts of the clade were found were not adjacent, then the clade names were numbered. Branches with hash marks are shown at one half their actual lengths. Euglenozoan OTUs are highlighted with blue.

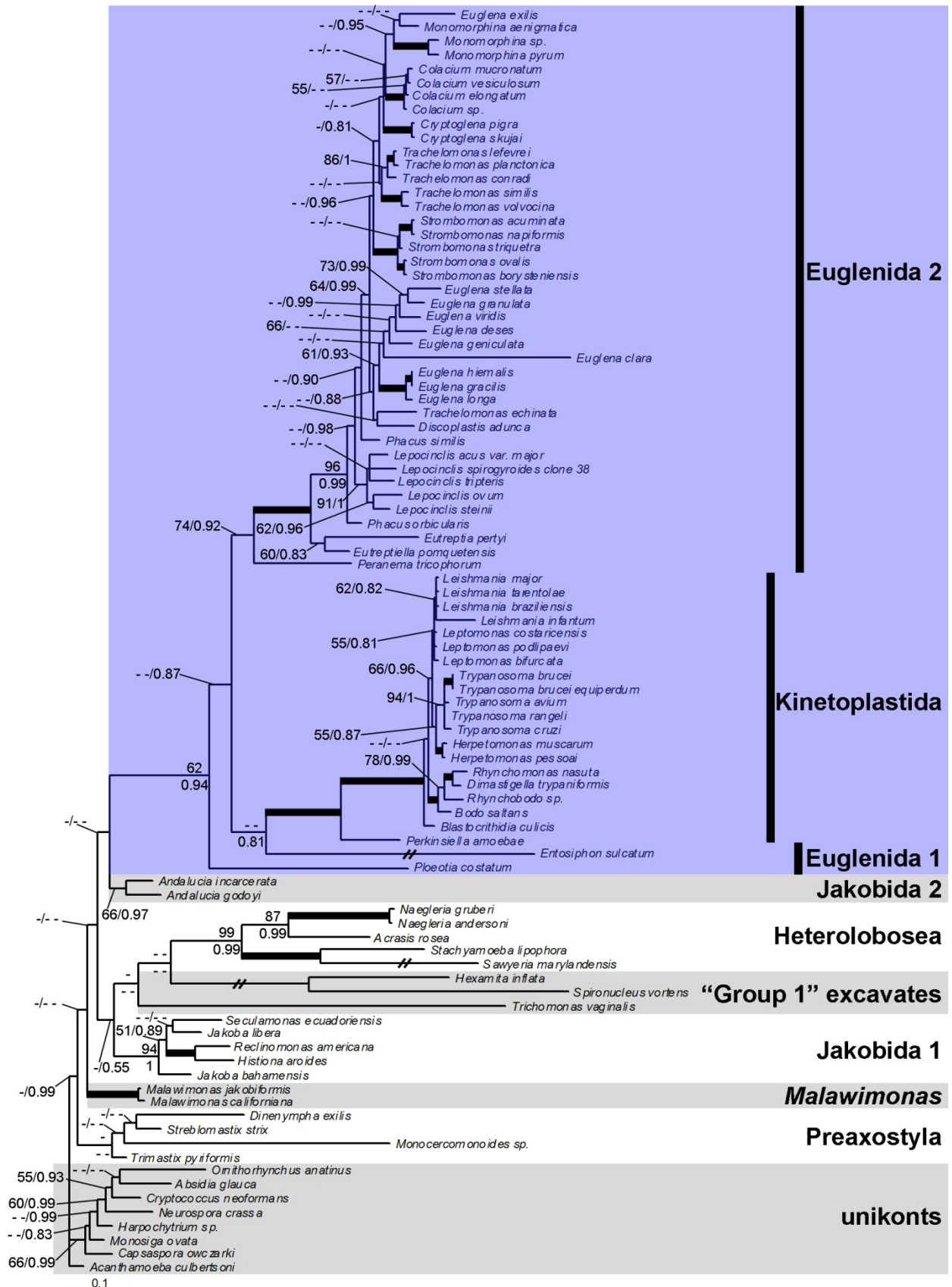
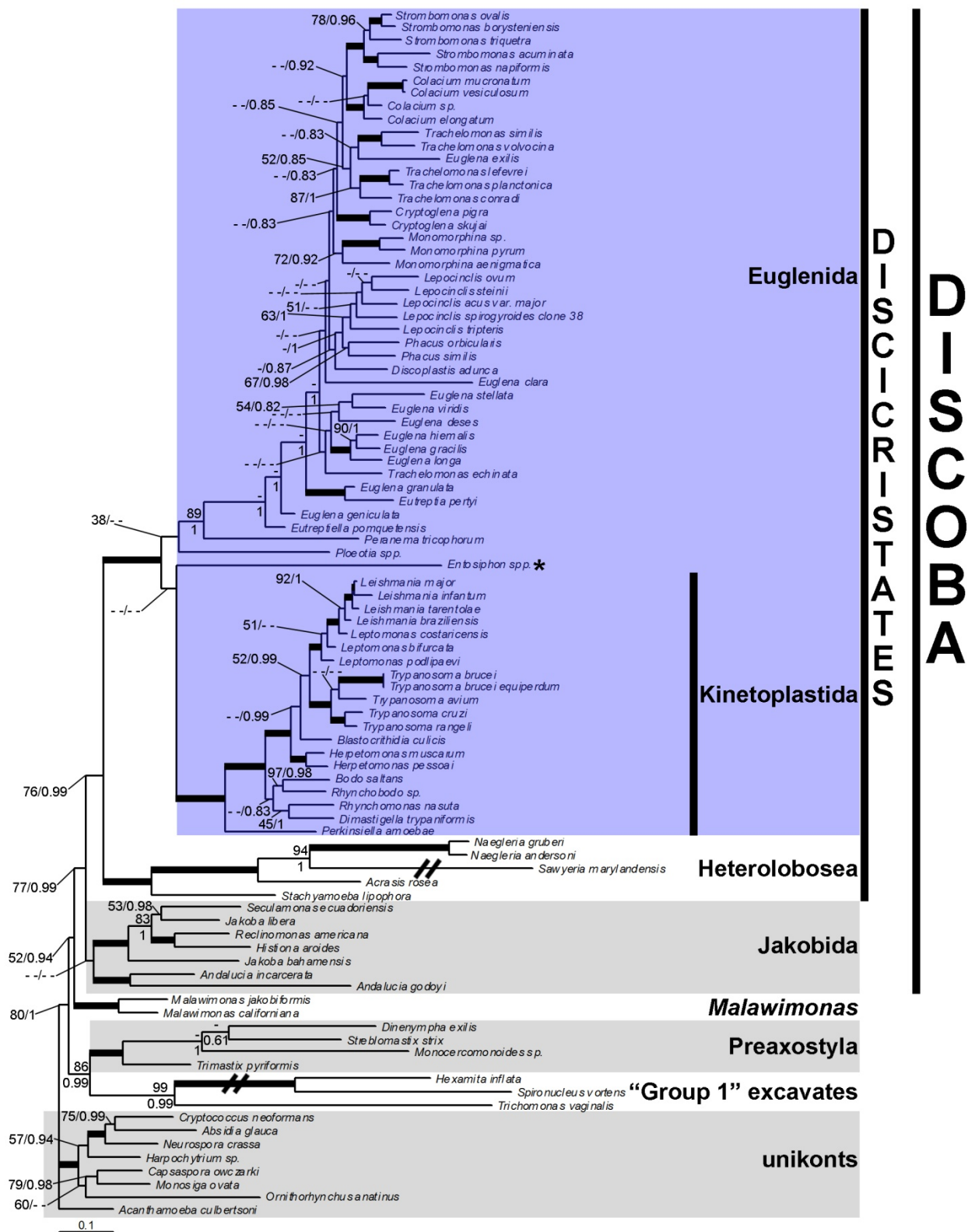


Figure 2.9. MrBayes consensus tree from the 1a18S-92 alignment analysis. The left or top value is the RAxML bootstrap score (bs) and the MrBayes posterior probability (pp) is below or to the right. Only $bs \geq 50\%$ and $pp \geq 0.80$ are shown here and the dark bars represent a node with $\geq 95\%$ bs from RAxML and 1.00 pp from the MrBayes resulting tree topologies for the same alignment. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Clade names are to the right with the deep level clades farther to the right and the more tip level clades closer to the OTUs. Branches with hash marks are shown at one half their actual lengths. Euglenozoan OTUs are highlighted with blue. The star denotes the OTU that was recovered as sister to *Ploeotia* spp. and basal in the recovered monophyletic Euglenida clade in the 1a3rd18S-84 MrBayes MLE alignment although it was not recovered within the Euglenida in this analysis.



CHAPTER 3

EUGLENID MULTIGENE PHYLOGENIES WITH EF1 α , 18S AND 28S**Abstract:**

Euglenids (Euglenozoa/Excavata) are a diverse group of organisms that are represented by autotrophs, heterotrophs and osmotrophs. The morphological synapomorphy of this group are the interesting cytoskeletal modifications known as pellicle strips. The systematics of the euglenids is still currently being revised based on both morphological (e.g., pyrenoid characteristics, mucilaginous ornamentation, etc.) and molecular lines of evidence. The vast majority of molecular evidence has been derived from ribosomal genes; nuclear small subunit ribosomal DNA (18S), nuclear large subunit ribosomal DNA (28S), plastid small subunit ribosomal DNA (16S), or plastid large subunit ribosomal DNA (23S). There has been a recent effort to incorporate protein-coding genes into uncovering the phylogenetic history of the group because these genes can be more conserved and less prone to long branch attraction artifacts (LBA) relative to ribosomal genes. Here we incorporate evidence from the highly conserved, nuclear-encoded, protein-coding gene elongation factor 1 alpha (EF1 α) that has been used effectively to resolve relationships of organisms at this level. The single gene trees were used to detect incongruence of data sets. Generally, EF1 α could not resolve certain clades due to its conservation and other aspects of its possibly complex evolutionary history within the Euglenida and the 18S and 28S genes were more prone to LBA. The concatenated genes did help to resolve and recover relationships within the euglenids, and the inclusion of EF1 α improved the

support of clades when compared to the nearly identical topology of the concatenation involving only the ribosomal genes. However, it appears that there are multiple copies of the gene or gene editing from gDNA to mRNA of EF1 α in the Euglenida that might interfere with the resolution of that group.

Introduction:

The euglenids contain photosynthetic, heterotrophic and osmotrophic species (Leedale 1967; Triemer, Farmer 2007). The synapomorphy of the euglenids is the pellicle, which is the complex containing of the plasma membrane covering proteinaceous strips that are arranged in parallel as well as the microtubules and endoplasmic reticulum that support them (Sommer 1965; Leander, Farmer 2000; Leander, Farmer 2001; Leander, Witek, Farmer 2001). The Euglenida are of interest for several reasons: they are ecologically important organisms (Xavier, Mainardespinto, Takino 1991), they are closely related to important human pathogens responsible for Leishmaniasis, African Sleeping Sickness and Chagas disease (Lukes et al. 2002), and they are part of the supergroup Excavata which is believed to be the earliest diverging lineage of the Eukaryotic branch of the tree of life (ToL) (Baldauf 2003; Burki, Pawlowski 2006; Burki, Shalchian-Tabrizi, Pawlowski 2008).

Previous clades and nomenclature of euglenids were based primarily on morphological structures that were visible with light microscopy such as gross marginal plastid morphology, pellicle rigidity or ability to drastically change shape (euglenoid metaboly), ability to form colonies with cells attached to one another via mucilaginous stalks, or the presence of a hardened mucilaginous case (lorica) (Leedale 1967). However, many of these characters were discovered to be plastic traits and did not correspond with the true species relationships within the group (Leander, Farmer 2001; Brosnan et al. 2003; Brown, Zakrys, Farmer 2003; von der Heyden et al.

2004; Triemer et al. 2006; Ciugulea et al. 2007; Ciugulea et al. 2008). Many of the currently known genera were actually polyphyletic and the euglenids required a major revision.

The major revisions that aided in supporting “true species phylogenies” of euglenids were recovered from a number of lines of evidence including morphological characters detectable only with electron microscopy (EM) and molecular data. The pellicles of the euglenids have a number of characters that can be used to reconstruct the phylogeny of the the group and they can also be mapped onto gene trees to give ideas of when certain morphological changes occurred in the evolution of the euglenid pellicle (Leander, Farmer 2000; Leander, Farmer 2001; Leander, Witek, Farmer 2001). Some of these phylogenetically informative characters include: pellicle pores and their distribution along the pellicle, pellicle strip reduction events at the posterior and anterior ends of the cells, pellicle strip projections, frame shapes of the pellicle strips.

Plastid morphological characters that determined relationships in original classification in the euglenids were not accurate phylogenetic markers, although recently discovered characters that are useful in recovering the Euglenida. Historically, only gross morphological margins, the presence or absence of pyrenoids (a concentrated region of RuBisCO within the plastid), and the presence or absence of paramylon (a beta-1,3 glucan common to all euglenids) were the only plastid characters used to classify euglenids. However, these were found to lead to erroneous classification of species because some of these items are ephemeral or differentially characterized in the cell cycle stage of the unicellular euglenids. Brown, Zakrys and Farmer (2003) uncovered the current understanding of the reduced plastid diversity and possible character states for use in phylogeny reconstruction. These characters included pyrenoid morphology and position as well as paramylon grain(s) number and position with respect to the euglenid plastid(s).

A deeper-level clade that remains controversial is the “loricate clade” which contains two genera which have the ability to secrete mucilage and harden it by incorporating minerals, *Trachelomonas* and *Strombomonas*. Originally, these two genera were both within *Trachelomonas* until Deflandre (1930) split them based on lorica characteristics specific to each clade. Recently, Marin et al. (2003) merged the two back together, however, Brosnan et al. (2005) helped clarify this using differences between the two genera with respect to lorica development and pellicle strip morphology to support the current genera. Brosnan et al. (2005) have also hypothesized the close relationship between the genus *Colacium*, whose members secrete mucilaginous stalks from their flagellar canal, and the “loricate clade” to form a “mucilaginous clade”. The hypothetical synapomorphy of this clade is the mass production of mucilage as compared to other euglenids which only have minimal mucilage production.

Unfortunately, most remaining resolution has only been based on ribosomal genes such as nuclear small subunit ribosomal DNA (18S), nuclear large subunit ribosomal DNA (28S), plastid small subunit ribosomal DNA (16S) or plastid large subunit ribosomal DNA (23S) (Zakrys 1988; Linton et al. 2000; Brosnan et al. 2003; von der Heyden et al. 2004; Milanowski et al. 2006; Triemer et al. 2006; Ciugulea et al. 2007; Kosmala et al. 2007; Kim, Shin 2008). Allegedly, there is little phylogenetic signal left in this 18S molecular marker (Busse, Preisfeld 2002; Busse, Preisfeld 2003). Also, 18S is relatively short with fewer phylogenetically informative sites and is prone to long branch attraction (LBA) artifacts, which lead to strongly supported yet incorrect relationships of taxa, although this has been less studied when using the other ribosomal genes. The relationships within the Euglenozoa could be robustly resolved if appropriate novel markers with different evolutionary constraints than the ribosomal genes are used. These markers would provide different lines of evidence that may increase overall

underlying yet hidden support that ribosomal genes alone may not recover when all genes are analyzed in concatenation (de Queiroz, Gatesy 2007). There has been a recent effort to incorporate protein-coding genes into uncovering the phylogenetic history of the group because those genes tend to be more conserved and less prone to LBA as compared to the ribosomal genes (Breglia, Slamovits, Leander 2007); however, it has proven problematic because of complex gene families (Simpson, Roger 2004) and gene loss or horizontal gene transfer (HGT) as reported in Chapters 1 and 2.

However, in this study we focus on recovering the phylogeny of the euglenids by introducing evidence from the highly conserved, nuclear-encoded, protein-coding gene elongation factor 1 alpha ($EF1\alpha$) that has been used effectively to determine relationships of organisms at this level (Baldauf, Doolittle 1997; Tanabe et al. 2004; Fiore-Donno et al. 2005). The single gene trees were used to detect incongruence of data sets based on comparisons with the analysis and resulting euglenid phylogeny of Triemer et al. (2006) based on the shared organisms in both studies. Generally, the shortcomings of the genes were that $EF1\alpha$ could not resolve certain clades due to its conservation and other aspects of its possibly complex evolutionary history within the Euglenida and the 18S and 28S genes were more prone to LBA. The concatenated gene phylogenies did help to resolve and recover relationships within the Euglenida, but with several important differences from previous studies involving only ribosomal genes (Brosnan et al. 2003; Triemer et al. 2006; Ciugulea et al. 2007; Ciugulea et al. 2008). However, it appears that there are complex evolutionary processes pertaining to $EF1\alpha$ within the euglenids such as multiple instances of gene duplication in various lineages and/or gene editing from gDNA to mRNA and this might be affecting the recovery of the true species tree of the Euglenida in this study.

Materials and Methods:

Culture information and nucleic acids extractions

All taxa with sources and accession numbers used in this study are shown in Table 3.1. Taxa were either obtained from culture collections or from colleagues (Triemer et al. 2006). All isolates were grown in the conditions and with the media specified in Chapter 1 of this thesis.

Sequence retrieval, nucleotide isolation, amplification, and sequencing

EF1 α , 18S and 28S genes were obtained from euglenids for ingroup taxa as well as kinetoplastids (basal heterotrophic bodonids as well as derived parasitic trypanosomatids) as outgroup taxa. This was done either with PCR or searches of EST and genomic databases (TBestDB, TriTrypDB v.1.1) or GenBank release 172 (Table 3.1). New partial 28S sequences were deposited in GenBank with accession numbers XXXX-XXXX: euglenid (3) and kinetoplastid (1).

Total genomic DNA and/or total RNA was extracted from uniprotistan cultures and amplified and sequenced as specified in Chapter 1 of this thesis. The 28S primers and PCR amplification protocols used were based on previously published research (Brosnan et al. 2003; Triemer et al. 2006).

Sequence Alignment

All EF1 α nucleotide sequences, representing 55 operational taxonomic units (OTUs), were aligned with MUSCLE (Edgar 2004) and then edited by eye against closely related organisms' EF1 α sequences based on amino acid conceptual translations in Geneious Pro v3.8.5 (Biomatters Ltd.; available through www.geneious.com) (Drummond et al. 2009) and previously published EF1 α alignments (Keeling, Inagaki 2004; Gile, Keeling 2008).

The 18S gene was acquired for the 55 OTUs represented in the EF1 α alignment and they were combined into a FASTA-formatted document which was aligned with SINA (SILVA Incremental Aligner) on the Silva release 100 comprehensive ribosomal RNA databases website (www.arb-silva.de/) to more accurately align the sequences based on the secondary structure database (Pruesse et al. 2007). This process was repeated for the 28S genes and OTUs and resulting alignment, however, *Strombomonas napiformis* and *Euglena clara* were missing sequence between their 28S contigs for the 5' and 3' ends of the gene.

The two sequences representing each respective species were forced together to create a single sequence for each species in the final alignment. Both ends of the 28S gene, for *S. napiformis* and *E. clara*, were included in the alignment as separate OTUs with all other organisms (57 OTUs included at this stage) and edited by eye as separate sequences. Question marks were placed after the 5' end of 28S for *S. napiformis* to match the gap, based on the alignment of the species in the alignment, between that sequence and the 3' end of the gene for the same organism. The 3' end OTU of *S. napiformis* was then copied and pasted to the end of the question marks to complete one single comprehensive sequence to represent *S. napiformis* in all analyses involving 28S. Finally, the short, separate *S. napiformis* 3' end OTU sequence was deleted from the alignment (56 OTUs at this stage) so that that portion of the gene for *S. napiformis* was not represented twice in the phylogenetic analyses. The same process was repeated for the 5' and 3' 28S sequences of *E. clara*. Therefore, the same 55 OTUs were represented in the final alignment of 28S such that it could be used in concatenation with the 55 OTU alignments of EF1 α and 18S.

All resulting gapped alignments from EF1 α , 18S and 28S were run through GBlocks v0.91b (Castresana 2000; Talavera, Castresana 2007) on the online server

(http://molevol.cmima.csic.es/castresana/Gblocks_server.html) to create reproducible alignments with unsupported residues eliminated. The GBlocks parameters selected were as follows: codon input data, allowed for smaller blocks, allowed for gap positions within final blocks and allowed for less strict flanking positions. The EF1 α alignment is Fig. 1.1 in Chapter 1, and the 18S and 28S alignments are Figs. 3.1 and 3.2, respectively. Third codon nucleotides were removed from the original EF1 α alignment to eliminate any homoplasious noise associated with that position to create new alignment 1a3rd. Concatenation alignments of EF1 α , 18S, and 28S were created such that the GBlocks alignments for each were put end to end, resulting in 18S28S, 1a18S28S and 1a3rd18S28S. All details for the alignments are in Table 3.2.

Phylogenetic analysis

All alignments were processed with Maximum likelihood (ML) and Maximum likelihood estimation using Bayesian algorithms (MLE). ML was performed using RAxML v7.0.4 with simultaneous autodetect rapid bootstrapping (Stamatakis 2006; Stamatakis, Hoover, Rougemont 2008) on the CIPRES server (Miller et al. 2009). MLE was also performed with MrBayes v3.1.2 (Huelsenbeck, Ronquist 2001; Ronquist, Huelsenbeck 2003) on either the CIPRES server or the University of Georgia (UGA) rCluster. Effects of the question marks within the *S. napiformis* and *E. clara* sequences in the alignments involving 28S were minimized by scoring them as missing data and not gaps during phylogenetic analyses. For both types of analyses, the model of sequence evolution (MoSE) was the one specified from the AIC indices from MODELTEST v3.7 (Posada, Crandall 1998; Posada, Crandall 2001; Posada, Buckley 2004). Trees were rooted with kinetoplastids using a mixture of trypanosomatids and bodonids in general and *Rhynchomonas nasuta*, specifically. See Table 3.3 for further details of analyses and MoSE selection for each analysis.

Clade Recovery

Due to the controversy regarding the true species tree topology of extant euglenids, we compared our results to the previous clades recovered by Triemer et al. (2006) (Fig. 3.3 and summarized in Fig. 3.4A). We chose this study as a reference because it has the most overlap with our species sampling, and the largest number of characters sampled from at least two genes, in this case 18S and 28S, both of which were included in this alignment. Most other studies have a more limited taxon sampling or only use one ribosomal gene (Brosnan et al. 2003; Milanowski et al. 2006; Ciugulea et al. 2007; Kosmala et al. 2007; Ciugulea et al. 2008; Kim, Shin 2008; Kosmala et al. 2009). All alignments were considered using all analysis types and recovery of expected monophyletic clades was summarized (Table 3.4). The total number of recovered clades (out of the possible clades) if the gene trees match the species tree, were tabulated for each alignment type to analyze which methods were best at clade recovery. The alignments were expected to recover 26 clades based on the shared OTU sampling between this study and Triemer et al. (2006).

Results:

Single gene analyses

EF1 α is very conserved and generally results in a great deal of polytomies in the single gene tree; however, there is also evidence of recovery of clades that are not supported by any previous molecular or morphological phylogenetic analyses. In general, EF1 α does not recover expected clades, especially when the third codon position is removed (1a3rd; Fig. 3.4C) because much information present in the “all-nucleotide alignment” (1a; Fig. 3.4B) is lost and tip-level clades, such as the H3 clade of the subclade of the *Euglena* lineage, *Cryptoglena* and the N2 clade of the *Trachelomonas* lineage, are no longer recovered (Table 3.4; Figs. 3.5 and 3.6). No

genera are ever recovered with either the 1a or 1a3rd analyses. An unexpected result of the EF1 α single gene analyses is that the previously supported basal Euglenea (all euglenids derived from a phototrophic ancestor, otherwise referred to as the euglenophytes) clade known as the Eutreptiales was not recovered, although the two species represented in this study by that clade, *Eutreptia pertyi* and *Eutreptiella pomquetensis*, were recovered near one another but were supported as more closely related to members of the genus *Euglena* than to one another.

The 18S single gene phylogeny recovers the largest number of expected clades of any single-gene phylogeny (58%); however, it seems to be negatively affected by LBA due to certain euglenids with highly derived 18S sequences (Figs. 3.4D and 3.7). Both deep-level clades, such as clade “B” (all euglenophytes except the Eutreptiales and *Discoplastis*) and clade “C” (*Phacus* joined with *Lepocinclis*) as well as tip-level clades such as the genera *Phacus*, *Lepocinclis*, *Cryptoglena*, *Monomorphina*, *Colacium* and *Strombomonas* are recovered. However, some of the more erroneous relationships recovered in this tree were due to the long-branching species such as *Euglena exilis* and *E. clara*. The clade “I” representative species *Euglena exilis* is recovered as sister to the *Monomorphina* clade rather than sister to *Euglena* clade “H”. *Euglena clara* is recovered within the *Euglena* clade “H” species, but not within the “H3” clade where it was expected, based on Triemer et al. (2006).

The 28S single gene analyses were also slightly skewed due to LBA artifacts and unresolved relationships from polytomies, but this might be partly due to incomplete 28S sequences for certain organisms. In particular it is noted that the branches of the partial sequences for *S. napiformis*, *E. clara*, *Peranema trichophorum* and *Ploeotia oblonga* are extremely long although they seem to cluster with relative consistency in the portions of the tree where they were expected. *S. napiformis* was recovered within the *Strombomonas* clade and *P.*

trichophorum and *P. oblonga* clustered together to recover the Heteronematina clade basal to the Euglenida. As in the 18S phylogenies, *E. clara* among clade “H” *Euglena* species, but it did not fall in the “H3” clade.

Certain genera are recovered, such as *Colacium*, *Monomorphina*, *Cryptoglena*, and *Strombomonas*; however, there is little resolution at the deeper levels of the Euglenida, unless it is with clades that are not present in Triemer et al. (2006). For instance, *Colacium* grouped, albeit with weak support, as sister to a cluster of organisms including all the members of clade “C” (*Lepocinclis* joined with *Phacus*), although neither *Phacus*, *Lepocinclis* nor clade “C” were recovered as monophyletic. The exception was that *Cryptoglena* and *Monomorphina* were recovered together in clade “L” as expected. However, clade “L” was sister to *Trachelomonas echinata* in the 28S phylogenies rather than to the highly mucilaginous OTUs that comprise clade “K” (*Colacium* sister to the loricate clade “M”, which combines *Strombomonas* with *Trachelomonas*) that Triemer et al. (2006) recovered.

Concatenated gene analyses

Overall, the concatenated gene analyses were the best at recovering the previously reported relationships from Triemer et al. (2006) (Figs. 3.3, 3.9, 3.10 and 3.1). Both three-gene concatenated analyses recovered 61% of the expected clades (Table 3.4) with the ribosomal concatenated analyses recovering just less than that (58%), although there were some differences between the clades recovered depending upon analysis type or whether or not EF1 α was included or the third codon position of EF1 α was removed. For example, the MLE analysis of the concatenated gene analysis with the entire EF1 α gene (1a18S28S; Fig. 3.10) did not recover clade “J”, that Triemer et al. (2006) found was comprised of the members of the mucilaginous clade “K” and clade “L” (*Cryptoglena* sister to *Monomorphina*; Fig. 3.3), because the

problematic *Euglena exilis* OTU branched as sister to clade “L”. However, in the ML analysis of 1a18S28S (topology not shown) a novel clade comprised of *Discoplastis adunca* and clade “C” (*Phacus* joined with *Lepocinclis*) was sister to the *E. exilis* sister to clade “L” relationship that MLE recovered. Since all other aspects of these analyses were the same and the only difference between them was the use of ML or MLE, therefore, the analysis-type was responsible for the differences in euglenid relationships. However, when the third codon position is removed from the three-gene concatenated alignment (1a3rd18S28S) or when the EF1 α gene is removed completely (18S28S) both ML and MLE analyses recover clade “J” (Figs. 3.9 and 3.11).

There were several interesting differences between the clade recovery of this study and Figure 1 from Triemer et al. (2006) (Fig. 3.3). One clade that was consistently recovered was the grouping of the mucilaginous clade “K”; however, the loricate clade “M” (*Strombomonas* sister to *Trachelomonas*) was not recovered because *Strombomonas* was sister to *Colacium* rather than with the other loricate-forming genus, *Trachelomonas* (Figs. 3.9, 3.10 and 3.11). Also, the placement of *E. exilis* seems to be very difficult to resolve even through the use of gene concatenation as it is sometimes found branching sister to clade “L” (*Cryptoglena* sister to *Monomorphina*; Fig. 3.10) or within the *Discoplastis adunca* + clade “C” (*Phacus* sister to *Lepocinclis*) group (Figs. 3.9 and 3.11).

The latter clade is supported as being basal in the Triemer et al. (2006) phylogeny (Fig. 3.3), branching just after the Eutreptiales; however, in this study although the clade was typically recovered, the placement is sometimes found to be more recently diverged in the euglenid lineage as sister to clade “J” (ML 18S28S, MLE 1a18S28S and 1a3rd18S28S; Figs. 3.9, 3.10, 3.11) or sister to clade “L” (ML of 1a18S28S; topology not shown) although it is sometimes recovered where it is expected (ML of 1a3rd18S28S; topology not shown). These relationships

are completely unresolved in the 18S28S MLE analysis with lack of resolution of clades noted on Fig. 3.9, but the actual topology is not shown.

It is also interesting to note that although analyses on the 18S28S alignment were performed in this study, which were comparable to the Triemer et al. (2006) study except for differences in alignments and OTU sampling, the resultant topologies (Figs. 3.4F and 3.9) were in greater accordance with those of the 1a3rd18S28S (Figs. 3.4F and 3.11) followed closely by the 1a18S28S (Figs. 3.4F and 3.10) analyses from this study rather than the Triemer et al. (2006) topology (Figs. 3.3 and 3.4A). The 18S28S (ML; Fig. 3.9) and 1a3rd18S28S (ML and MLE; Fig. 3.11) topologies were identical, except that the overall trend was that the latter had increased bootstrap (bs) and posterior probability (pp) for ML and MLE, respectively, when compared to the former. This occurred most often at the deeper level nodes (e.g., Euglenea, clades “L”, “K”, the relationship that placed *Strombomonas* sister to *Colacium*, and *Trachelomonas*). The 18S28S MLE topology (not shown) was identical to the 18S28S ML topology except that it was not able to resolve the relationships between clade “J”, clade “H”, *Lepocinclis*, the two *Phacus* species, *Discoplastis adunca* or *Euglena exilis* (Fig. 3.9), such that the majority of the Euglenea deep-level clades were not resolved.

Discussion:

Concatenation of genes helps to overcome single-gene deficiencies

Although single genes on their own tend to be affected by evolutionary mechanisms that disrupt the recovery of the phylogenetic signal that leads to the discovery of the true species tree of an organismal group, concatenating genes tends to help in overcoming these erroneous relationships. This was evidenced by our results in that the analyses that recovered the largest number of previously observed clades from Triemer et al. (2006) were the concatenated gene

analyses. Although there were slight differences in the clades recovered between the three concatenated gene alignments and their analysis types, they generally gave highly supported clades such as monophyletic genera and deeper-level relationships as well as tip-level relationships within genera. However, the inability for some of the deeper clades to reach a consensus on a congruent topology is troublesome and might be evidence that there are some molecular evolutionary mechanisms that cannot be overcome by gene concatenation alone for this data set.

Incongruity of recovered Euglenida clades compared to findings of Triemer et al. (2006)

Alignment differences between the Triemer et al. (2006) study and this study could have led to the differences in the resultant topologies. The Triemer et al. (2006) study used alignments that were completely edited by eye based on the secondary structures of the ribosomal genes of *Euglena gracilis* (Kjer 1995; Wuyts et al. 2001). This study instead used the automated alignment of SINA, which aligns the OTUs based on the secondary structure of all known eukaryote 18S and 28S, respectively, for each separate alignment. Those analyses were then checked by eye and processed through GBlocks to remove unsupported residues using parameters stated in the methods section, such that the alignments created could be easily reproduced by other researchers. It is unclear how Triemer et al. (2006) defined the “conserved areas” that they included in their alignments or what the final alignment was since it was not published.

The Triemer et al. (2006) study had differential representation of characters from the 18S and 28S genes as compared to this study. Triemer et al. (2006) included more sites than this study for 18S (1807 versus 1324) and less sites than this study for 28S (903 versus 1452), which led to a longer final concatenation for this study (2776 versus 2711) that was more equivalent in

its ratio of representation by the number of residues from the two genes (18S:28S; ~2:1 versus ~1:1 in this study). However, some particular species in this study were used even though they had incomplete sequences with gaps that had to be scored as missing data (e.g., 28S for *Strombomonas napiformis* and *Euglena clara*). This could have resulted in the abnormally long branches of those OTUs for the 28S gene, and although they are recovered properly in their respective genera (in all analyses that included the 28S data set), they are not recovered with the same species-level relationships as they were in Triemer et al. (2006) (Figs. 3.3, 3.8, 3.9, 3.10, 3.11).

Another factor that could have led to the topology incongruence between studies involves taxon sampling. Triemer et al. (2006) had 84 taxa in their full analysis with two members of the eutreptiales, *Eutreptia viridis* and *Eutreptiella* sp. (CCMP 389), set as the outgroups to the rest of the Euglenia taxa. All 84 taxa are members of the Euglenia, meaning that they are all descended from a phototrophic euglenid ancestor that acquired its plastid through a secondary endosymbiotic event involving a green algal cell. This study only had 40 members of the Euglenia represented, and of those only 38 were common to both studies since *Strombomonas napiformis* and *Colacium* sp. (NJ) were not included in Triemer et al. (2006). Some clades from Triemer et al. (2006) were only represented by one species (e.g., *Discoplastis adunca* in clade “A”, *Euglena exilis* in clade “I”, and *Trachelomonas echinata* in clade “N1”) or not at all (e.g., clade “E2” within *Phacus*). This could have contributed to the lack of resolution or differential resolution of clades at deeper levels (e.g., 18S28S MLE topology and 18S28S ML, respectively; Fig. 3.9) as compared to Figure 1 from Triemer et al. (2006) (Fig. 3.3).

Although this study had less depth of coverage of the photosynthetic Euglenia, it increased breadth of sampling by including colorless euglenids, the Heteronematina, as well as

kinetoplastids as the outgroup taxa. The outgroup taxa included a mixture of the both the derived parasitic trypanosomatids as well as the heterotrophic bodonids, which are thought to retain more characteristics of the ancestral kinetoplastids. Therefore a member of the bodonids, *Rhynchomonas nasuta*, was forced as the root for this study. The inclusion of the kinetoplastids as an outgroup allowed the evaluation of the recovery of the Euglenida and Euglenea clades and whether or not the Heteronematina were basal to the Euglenea and whether or not the eutreptiales, in this case *Eutreptia pertyi* and *Eutreptiella pomquetensis*, could be supported as the most basal members of the Euglenea without being forced as in Triemer et al. (2006).

All of these differences between the two studies combine to provide possible explanations as to the discrepancies between the topologies. The differential recovery of *Euglena exilis* (placement varied depending upon which analysis type and alignment was used) in this study does not appear to be due to any error in the alignment process when checked *ad hoc*, but it is more likely due to the character sampling and editing in the alignment, as well as both ingroup and outgroup taxon sampling. When the entire *E. exilis* 28S sequence is BLASTed on GenBank its first match is to itself, but then it seems to share an insertion that members of *Phacus* have which could have led to the sister relationship of *E. exilis* with *Phacus orbicularis* in the 28S single-gene analyses (Fig. 3.8) and the 18S28S (ML) and 1a3rd18S28S concatenated gene analyses (Fig. 3.9 and 3.11). This could have also contributed to the lack of resolution in the 18S28S MLE topology (not shown).

However, it is also possible that the signal of *E. exilis* and the OTUs it is tends to be closely related to in the 18S analyses compared to those in the 28S analyses are in direct conflict with each other such that when the ribosomal genes were concatenated (18S28S) this resulted in the large polytomy in which those OTUs were involved. When the same BLAST analysis is

performed with the complete *E. exilis* 18S sequence there is also a variable inserted portion that seems similar to some members of *Euglena*, but members of *Monomorphina*, *Phacus* and *Lepocinclis* also seem to share the insertion at the same position. This could lead to the support of the relationship between *E. exilis* and either *Monomorphina* using 18S in the single-gene analysis (Fig. 3.7), with clade “L” (*Monomorphina* sister to *Cryptoglena*) in the 1a18S28S concatenated analysis (Fig. 3.10), or with clade “C” (*Phacus* sister to *Lepocinclis*) as described earlier in reference to the 28S sequence of *E. exilis*.

3rd codon positions and possible incongruity of EF1 α in recovering Euglenida clades

The topologies that resulted from the EF1 α single gene analyses were incongruent with one another and with the 18S and 28S single gene phylogenies as well as the concatenated gene phylogenies. Although the third codon position is sometimes problematic in terms of being saturated with mutations due to the degeneracy of the amino acid code, it appears that in the single gene analyses the third codon position actually aids in recovering a more resolved clade with groups of organisms clustered together that are seen in other molecular analyses and are supported by morphological synapomorphies as described in Triemer et al. (2006). Therefore, we hypothesize that there are various evolutionary mechanisms within the EF1 α history in the Euglenida that are as yet undefined but that appear to disagree with the currently understood tip and deep-level relationships within this group of organisms.

Conclusion:

Although the evolutionary histories of the ribosomal genes are very different from EF1 α history within the Euglenida, this study found that the concatenation of all lines of evidence provided the greatest resolution of the group. Each gene type had problems. The ribosomal genes were affected by LBA and missing data, whereas the molecular evolution of EF1 α in the

euglenids was more complicated than anticipated. The latter was evidenced in that the eutreptiales were not recovered as a clade, nor were the representatives of the eutreptiales clade recovered as basal to the Euglenea as previously reported in Triemer et al. (2006). Also, none of the genera were recovered. In future studies, the concatenated gene analyses could be analyzed using partitioned Bremer support (PBS; Bremer 1994; Lambkin 2004) to determine the support or conflict at each node that is provided by each gene partition to help uncover the “true” phylogenetic relationships based on signal of the concatenated data set as opposed to the noise. This differential signal of data sets seems to be especially evident in the case of *Euglena exilis* since the concatenation of the ribosomal data sets (18S/28S) resulted in a lack of resolution at deeper levels and a large polytomy including *E. exilis*, whose placement was resolved, albeit with different species-level relationships depending on the data set used, in the single-gene analyses of 18S and 28S as well as the concatenated analyses that included EF1 α .

The incongruencies between the Euglenida “species tree” based on results from previous findings (Leander, Farmer 2001; Brosnan et al. 2003; Brown, Zakrys, Farmer 2003; von der Heyden et al. 2004; Triemer et al. 2006; Ciugulea et al. 2007; Ciugulea et al. 2008) and those of the Euglenida EF1 α “gene tree” from this study could be due to a complex evolutionary history of EF1 α within the Euglenida. Some possible causes for the disparity of the Euglenida EF1 α “gene tree” could be due to gene duplication events (followed by differential gene loss), HGT, other history-eradicating molecular mechanisms, or some combination of these phenomena.

Further examination of the molecular evolution of EF1 α in the Euglenida lineage should be explored further to clarify this incongruity with the previous phylogenetic relationships of the euglenids based on previous molecular and morphological studies (Leander, Farmer 2000; Leander, Witek, Farmer 2001; Brosnan et al. 2003; Brown, Zakrys, Farmer 2003; Brosnan et al.

2005; Milanowski et al. 2006; Triemer et al. 2006; Ciugulea et al. 2007; Ciugulea et al. 2008; Kim, Shin 2008; Kosmala et al. 2009). This should involve further sequencing of PCR products as well as Southern blots of EF1 α in euglenid species to determine copy number of that gene within taxa that were problematic in this study (e.g., eutreptiales, *Euglena* species, *Strombomonas* species, etc.). Also, sequencing of genomic DNA (gDNA) as well as mRNA of EF1 α should be pursued since there was evidence in Chapter 1 and 2 that the gDNA of *Ploeotia oblonga* was basal to the Euglenea, as previously observed in prior studies; however, the mRNA from the same species was recovered nested within the Euglenea along with the mRNA sequence of another species within the same genus, *P. costatum*. This indicates that there might be some informational change between gDNA and mRNA or that there are multiple copies of EF1 α within this lineage.

References:

- Baldauf, SL. 2003. The deep roots of eukaryotes. *Science* 300:1703-1706.
- Baldauf, SL, WF Doolittle. 1997. Origin and evolution of the slime molds (Mycetozoa). *Proceedings of the National Academy of Sciences of the United States of America* 94:12007-12012.
- Breglia, SA, CH Slamovits, BS Leander. 2007. Phylogeny of phagotrophic euglenids (Euglenozoa) as inferred from HSP90 gene sequences. *Journal of Eukaryotic Microbiology* 54:86-92.
- Brosnan, S, PJP Brown, MA Farmer, RE Triemer. 2005. Morphological separation of the euglenoid genera *Trachelomonas* and *Strombomonas* (Euglenophyta) based on lorica development and posterior strip reduction. *Journal of Phycology* 41:590-605.
- Brosnan, S, W Shin, KM Kjer, RE Triemer. 2003. Phylogeny of the photosynthetic euglenophytes inferred from the nuclear SSU and partial LSU rDNA. *International Journal of Systematic and Evolutionary Microbiology* 53:1175-1186.
- Brown, PJP, B Zakrys, MA Farmer. 2003. Plastid morphology, ultrastructure, and development in *Colacium* and the loricate euglenophytes (Euglenophyceae). *Journal of Phycology* 39:115-121.
- Burki, F, J Pawlowski. 2006. Monophyly of Rhizaria and multigene phylogeny of unicellular bikonts. *Molecular Biology and Evolution* 23:1922-1930.
- Burki, F, K Shalchian-Tabrizi, J Pawlowski. 2008. Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. *Biology Letters* 4:366-369.
- Busse, I, A Preisfeld. 2002. Phylogenetic position of *Rhynchopus* sp. and *Diplonema ambulator* as indicated by analyses of euglenozoan small subunit ribosomal DNA. *Gene* 284:83-91.
- Busse, I, A Preisfeld. 2003. Application of spectral analysis to examine phylogenetic signal among euglenid SSU rDNA data sets (Euglenozoa). *Organisms Diversity & Evolution* 3:1-12.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17:540-552.
- Ciugulea, I, MA Nudelman, S Brosnan, RE Triemer. 2007. Phylogeny of the euglenoid loricate genera *Trachelomonas* and *Strombomonas* (Euglenophyta) inferred from nuclear ssu and lsu rDNA. *Journal of Phycology* 43:196.
- Ciugulea, I, MA Nudelman, S Brosnan, RE Triemer. 2008. Phylogeny of the euglenoid loricate genera *Trachelomonas* and *Strombomonas* (Euglenophyta) inferred from nuclear SSU and LSU rDNA. *Journal of Phycology* 44:406-418.

- de Queiroz, A, J Gatesy. 2007. The supermatrix approach to systematics. *Trends in Ecology & Evolution* 22:34-41.
- Drummond, AJ, B Ashton, M Cheung, J Heled, M Kearse, R Moir, S Stones-Havas, T Thierer, A Wilson. 2009. Geneious: Biomatters Ltd.
- Edgar, RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792-1797.
- Fiore-Donno, AM, CD Berney, J Pawlowski, SL Baldauf. 2005. Higher-order phylogeny of plasmodial slime molds (Myxogastria) based on elongation factor 1-A and small subunit rRNA gene sequences. *Journal of Eukaryotic Microbiology* 52:201-210.
- Gile, GH, PJ Keeling. 2008. Nucleus-encoded periplastid-targeted EFL in chlorarachniophytes. *Molecular Biology and Evolution* 25:1967-1977.
- Huelsenbeck, JP, F Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.
- Keeling, PJ, Y Inagaki. 2004. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1 alpha. *Proceedings of the National Academy of Sciences of the United States of America* 101:15380-15385.
- Kim, JI, W Shin. 2008. Phylogeny of the euglenales inferred from plastid LSU rDNA sequences. *Journal of Phycology* 44:994-1000.
- Kjer, KM. 1995. Use of ribosomal-RNA secondary structure in phylogenetic studies to identify homologous positions - an example of alignment and data presentation from the frogs. *Molecular Phylogenetics and Evolution* 4:314-330.
- Kosmala, S, A Karnkowska-Ishikawa, R Milanowski, J Kwiatowski, B Zakrys. 2009. Phylogeny and systematics of *Euglena* (Euglenaceae) species with axial, stellate chloroplasts based on morphological and molecular data-new taxa, emended diagnoses and epitypifications. *Journal of Phycology* 45:464-481.
- Kosmala, S, R Milanowski, K Brzoska, G Pekala, J Kwiatowski, B Zakrys. 2007. Phylogeny and systematics of the genus *Monomorpha* (Euglenaceae) based on morphological and molecular data. *Journal of Phycology* 43:171-185.
- Leander, BS, MA Farmer. 2000. Comparative morphology of the euglenid pellicle. I. Patterns of strips and pores. *Journal of Eukaryotic Microbiology* 47:469-479.
- Leander, BS, MA Farmer. 2001. Evolution of Phacus (Euglenophyceae) as inferred from pellicle morphology and SSU rDNA. *Journal of Phycology* 37:143-159.
- Leander, BS, RP Witek, MA Farmer. 2001. Trends in the evolution of the euglenid pellicle. *Evolution* 55:2215-2235.

- Leedale, G. 1967. Euglenoid Flagellates. 1-73.
- Linton, EW, MA Nudelman, V Conforti, RE Triemer. 2000. A molecular analysis of the euglenophytes using SSU rDNA. *Journal of Phycology* 36:740-746.
- Lukes, J, DL Guilbride, J Votypka, A Zikova, R Benne, PT Englund. 2002. Kinetoplast DNA network: Evolution of an improbable structure. *Eukaryotic Cell* 1:495-502.
- Marin, B, A Palm, M Klingberg, M Melkonian. 2003. Phylogeny and taxonomic revision of plastid-containing euglenophytes based on SSU rDNA sequence comparisons and synapomorphic signatures in the SSU rRNA secondary structure. *Protist* 154:99-145.
- Milanowski, R, S Kosmala, B Zakrys, J Kwiatowski. 2006. Phylogeny of photosynthetic euglenophytes based on combined chloroplast and cytoplasmic SSU rDNA sequence analysis. *Journal of Phycology* 42:721-730.
- Miller, MA, MT Holder, R Vos, PE Midford, T Liebowitz, L Chan, P Hoover, T Warnow. 2009. The CIPRES Portals.
- Posada, D, TR Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53:793-808.
- Posada, D, KA Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817-818.
- Posada, D, KA Crandall. 2001. Selecting the best-fit model of nucleotide substitution. *Systematic Biology* 50:580-601.
- Pruesse, E, C Quast, K Knittel, BM Fuchs, W Ludwig, J Peplies, FO Glockner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucl. Acids Res.* 35:7188-7196.
- Ronquist, F, JP Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
- Simpson, AGB, AJ Roger. 2004. Protein phylogenies robustly resolve the deep-level relationships within Euglenozoa. *Molecular Phylogenetics and Evolution* 30:201-212.
- Sommer, JR. 1965. Ultrastructure of pellicle complex of *Euglena gracilis*. *Journal of Cell Biology* 24:253-&.
- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.
- Stamatakis, A, P Hoover, J Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Systematic Biology* 57:758-771.

- Talavera, G, J Castresana. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. 56:564-577.
- Tanabe, Y, M Saikawa, MM Watanabe, J Sugiyama. 2004. Molecular phylogeny of Zygomycota based on EF-1 alpha and RPB1 sequences: limitations and utility of alternative markers to rDNA. Molecular Phylogenetics and Evolution 30:438-449.
- Triemer, RE, MA Farmer. 2007. A decade of euglenoid molecular phylogenetics. In: J Brodie, J Lewis, editors. Unraveling the algae: the past, present and future of algal systematics: Systematics Association Series, CRC Press. p. 315-330.
- Triemer, RE, E Linton, W Shin, A Nudelman, A Monfils, M Bennett, S Brosnan. 2006. Phylogeny of the euglenales based upon combined SSU and LSU rDNA sequence comparisons and description of *Discoplastis* gen. nov (Euglenophyta). Journal of Phycology 42:731-740.
- von der Heyden, S, EE Chao, K Vickerman, T Cavalier-Smith. 2004. Ribosomal RNA phylogeny of bodonid and diplomonid flagellates and the evolution of euglenozoa. Journal of Eukaryotic Microbiology 51:402-416.
- Wuyts, J, P De Rijk, Y Van de Peer, T Winkelmans, R De Wachter. 2001. The European Large Subunit Ribosomal RNA Database. Nucleic Acids Research 29:175-177.
- Xavier, MB, CSR Mainardespinto, M Takino. 1991. *Euglena sanguinea* Ehrenberg bloom in a fish-breeding tank (Pindamonhangaba, Sao-Paulo, Brazil). Archiv Fur Hydrobiologie:133-142.
- Zakrys, B. 1988. The nuclear DNA level as a potential taxonomic character in *Euglena* Ehr. (Euglenophyceae). Arch. Hydrobiol. Suppl. 78:483-504.

Table 3.1. Organism information for sequences used in this study. Bold sequences illustrate new sequences obtained in this study, an asterisk denotes different strain used for one or more of the genes in the concatenation study, and “Triemer” denotes sequences received from Triemer lab (MSU) that have not yet been published on GenBank. “Group” classification information was taken from Adl et al. 2005.

“Group(s)”	Species	Strain	EF1α GenBank/TBest DB Accession #	18S GenBank/TBest DB Accession #	28S GenBank/TBest DB Accession #
Euglenozoa/Euglenida/Euglenea	<i>Colacium elongatum</i>	M11	apply for; cDNA	Triemer	Triemer
Euglenozoa/Euglenida/Euglenea	<i>Colacium mucronatum</i>	UTEX LB 2524	apply for; cDNA	AF326232	EF999906
Euglenozoa/Euglenida/Euglenea	<i>Colacium</i> sp.	NJ	apply for; cDNA	Triemer	DQ140118
Euglenozoa/Euglenida/Euglenea	<i>Colacium vesiculosum</i>	UW Lazienki	apply for; cDNA	DQ249874	EF999912
Euglenozoa/Euglenida/Euglenea	<i>Cryptoglena pigra</i>	CCAP 1212/1	apply for; cDNA	AJ532437	DQ140101
Euglenozoa/Euglenida/Euglenea	<i>Cryptoglena skujai</i>	SAG 10.88	apply for; cDNA	AJ532438	AY130236
Euglenozoa/Euglenida/Euglenea	<i>Discoplastis adunca</i>	ASW 08039	apply for; cDNA	AJ532453	DQ140102
Euglenozoa/Euglenida/Euglenea	<i>Euglena clara</i>	SAG 25.98	apply for; cDNA	AJ532423	DQ140114
Euglenozoa/Euglenida/Euglenea	<i>Euglena deses</i>	SAG 1224-23	apply for; cDNA	AJ532407	DQ140122
Euglenozoa/Euglenida/Euglenea	<i>Euglena exilis</i>	SAG 1224-11a	apply for; cDNA	DQ140147	DQ140103
Euglenozoa/Euglenida/Euglenea	<i>Euglena geniculata</i>	SAG 1224-21	apply for; cDNA	AY523034	AY523027
Euglenozoa/Euglenida/Euglenea	<i>Euglena gracilis</i>	Z strain	apply for; cDNA	AF283308	X53361
Euglenozoa/Euglenida/Euglenea	<i>Euglena granulata</i>	OR5	apply for; cDNA	AJ532422	DQ140099
Euglenozoa/Euglenida/Euglenea	<i>Euglena hiemalis</i>	CCAP 1224/35	apply for; cDNA	DQ140157	DQ140121
Euglenozoa/Euglenida/Euglenea	<i>Euglena</i> (aka <i>Astasia</i>) <i>longa</i>	UTEX 512?	FJ807237; mRNA	AJ532428	AY130223
Euglenozoa/Euglenida/Euglenea	<i>Euglena stellata</i>	IAM E-11 & UTEX 372*	apply for; cDNA	AF090868	AY130229*
Euglenozoa/Euglenida/Euglenea	<i>Euglena viridis</i>	NJ001 & ATCC PRA110*	apply for; cDNA	AY523036	AY523031*
Euglenozoa/Euglenida/Euglenea	<i>Eutreptia pertyi</i>	UTEX LB 1290	apply for; cDNA	AF081589	Triemer
Euglenozoa/Euglenida/Euglenea	<i>Eutreptiella pomquetensis</i>	CCMP 1491	apply for; cDNA	AJ532398	EU624012
Euglenozoa/Euglenida/Euglenea	<i>Lepocinclis acus</i> var. <i>major</i>	UTEX LB 1316	apply for; cDNA	AF152104	AY130226
Euglenozoa/Euglenida/Euglenea	<i>Lepocinclis ovum</i>	SAG 1244-8	apply for; cDNA	AF110419	AY130235
Euglenozoa/Euglenida/Euglenea	<i>Lepocinclis spirogyroides</i>	SAG 1224-10b (clone 38)	apply for; cDNA	AY935694	AY130228
Euglenozoa/Euglenida/Euglenea	<i>Lepocinclis steinii</i>	UTEX 523	apply for; cDNA	AF096993	AY130815
Euglenozoa/Euglenida/Euglenea	<i>Lepocinclis tripteris</i>	UTEX 1311	apply for; cDNA	AY935696	AY130230
Euglenozoa/Euglenida/Euglenea	<i>Monomorphina aenigmatica</i>	UTEX 1284	apply for; cDNA	AF190814	DQ140117
Euglenozoa/Euglenida/Euglenea	<i>Monomorphina pyrum</i>	UTEX 2354	apply for; cDNA	AF112874	AY130234
Euglenozoa/Euglenida/Euglenea	<i>Monomorphina</i> sp.	MI 73	apply for; cDNA	DQ140130	DQ140113
Euglenozoa/Euglenida/Euglenea	<i>Phacus orbicularis</i>	ACOI 996	apply for; cDNA	Triemer	DQ140126
Euglenozoa/Euglenida/Euglenea	<i>Phacus similis</i>	SAG 58.81	apply for; cDNA	AJ532467	AY130237
Euglenozoa/Euglenida/Euglenea	<i>Strombomonas acuminata</i>	SAG 1280-1, ACOI2467*, & S716**	apply for; cDNA	DQ249883*	AY359914

Euglenozoa/Euglenida/Euglenea	<i>Strombomonas borysteniensis</i>	NJ Stromb 10	apply for; cDNA	DQ140131	AY359920
Euglenozoa/Euglenida/Euglenea	<i>Strombomonas napiformis</i>	NJ Strombo 2	apply for; cDNA	apply for; cDNA & gDNA	apply for; cDNA & gDNA
Euglenozoa/Euglenida/Euglenea	<i>Strombomonas ovalis</i>	NJ S115	apply for; cDNA	DQ140133	AY359919
Euglenozoa/Euglenida/Euglenea	<i>Strombomonas triquetra</i>	NJ S116	apply for; cDNA	DQ140153	AY359917
Euglenozoa/Euglenida/Euglenea	<i>Trachelomonas conradi</i>	UTEX 1313	apply for; cDNA	AY015004	
Euglenozoa/Euglenida/Euglenea	<i>Trachelomonas echinata</i>	SAG 1283-22	apply for; cDNA	AY015001	
Euglenozoa/Euglenida/Euglenea	<i>Trachelomonas lefevrei</i>	SAG 1283-10	apply for; cDNA	DQ140136	AY359952
Euglenozoa/Euglenida/Euglenea	<i>Trachelomonas planctonica</i>	Woongghi Trach 1 (WT1)	apply for; cDNA	DQ140138	AY359954
Euglenozoa/Euglenida/Euglenea	<i>Trachelomonas similis</i>	SAG 1283-14	apply for; cDNA	DQ140142	AY359948
Euglenozoa/Euglenida/Euglenea	<i>Trachelomonas volvocina</i>	UTEX 1327	apply for; cDNA	AJ532451	AY359953
Euglenozoa/Euglenida/Heteronematina	<i>Peranema trichophorum</i>	CCAP 1260/1b	FJ807243; mRNA	AF386636	apply for; cDNA & gDNA
Euglenozoa/Euglenida/Heteronematina	<i>Ploeotia oblonga</i>	ATCC 50573 (* = <i>P. costatum</i>)	apply for; gDNA	AF525486*	apply for; gDNA
Euglenozoa/Kinetoplastida/Metakinoplastina/Eubodonida	<i>Bodo saltans</i>	CCAP 1907/2	FJ807252; gDNA	AY490233	EF681898
Euglenozoa/Kinetoplastida/Metakinoplastina/Neobodonida	<i>Dimastigella trypaniformis</i>	ATCC 50263	FJ807253; gDNA	AY028447	AY028447
Euglenozoa/Kinetoplastida/Metakinoplastina/Neobodonida	<i>Rhynchobodo</i> sp.	ATCC 50359	FJ807264; gDNA	U67183	AY028453
Euglenozoa/Kinetoplastida/Metakinoplastina/Neobodonida	<i>Rhynchomonas nasuta</i>	ATCC 50292	FJ807253; mRNA	AY425023	DQ086724
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Blastocrithidia culicis</i>	ATCC 30268	FJ807251; gDNA	U05679	L19404
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Herpetomonas muscarum</i>	ATCC 30260 (* = <i>H. roitmani</i>)	FJ807255; gDNA	L18872	AF038026*
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Leishmania braziliensis</i>	MHOM/BR/75/M2904	XM_001563727; mRNA	GQ332355	U73613
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Leishmania infantum</i>	JPCM5	XM_001464627; mRNA	AM502245	AM502245
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Leishmania major</i>	Friedlin	XM_001682206; mRNA	X53915	CP000079
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Leishmania tarentolae</i>	Parrot-II	apply for; gDNA	M84225	apply for; gDNA
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Leptomonas bifurcata</i>	(* = <i>Leptomonas</i> sp.)	FJ807258; gDNA	DQ910925	L19409*
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Trypanosoma brucei</i>	TREU927	XM_817372; mRNA	M12676	NC_008409
Euglenozoa/Kinetoplastida/Metakinoplastina/Trypanosomatida	<i>Trypanosoma cruzi</i>	CL Brener	XM_814346; mRNA	AF359495	L22334

Table 3.2. Summary of alignments (55 OTUs). Alignment short names are abbreviations of the gene(s) and type of positions included in the alignment as detailed in separate columns of the table. Genes are EF1 α =1a or EF1 α with third codon positions removed=1a3rd, small subunit ribosomal DNA=18S, large subunit ribosomal DNA=28S and the concatenated alignments utilizing more than one gene are either 18S28S, 1a18S28S, or 1a3rd18S28S. Alignments with the 3rd codon nucleotide positions removed in the EF1 α sequences are denoted with an “X”. Number of base pairs (bp) in each alignment is also stated.

Short Name	Gene(s)	3 rd positions removed	bp
1a	EF1 α		1277
1a3 rd	EF1 α	X	818
18S	18S		1324
28S	28S		1452
18S28S	18S & 28S		2776
1a18S28S	EF1 α 18S & 28S		4003
1a3 rd 18S28S	EF1 α 18S & 28S	X	3594

Table 3.3. Summary of tree statistics from empirical data and results from ML and MLE analyses. Analysis details and parameters are listed under the analysis type; MoSE=model of sequence evolution specified by Modeltest v3.7 used in ML and MLE for that alignment, #Gen=number of generations used in MrBayes; burn-in=relative burn in fraction value; #trees=number of trees used to calculate support. Asterisk (*) denotes that although four original runs of 30million generations, sampled every 1000 generations, with 0.5 burn-in were created, all the trees after burn-in from all four runs were combined to create one 50% majority rules consensus tree.

Analysis Type & Parameters	Alignment Input	MoSE	#Gen (in millions)	burn-in	# trees
<u>RAxML (ML)</u> autodetect bootstrap replicates	1a	GTR+I+ Γ	n/a	n/a	350
	1a3 rd	GTR+I+ Γ	n/a	n/a	550
	18S	GTR+ Γ	n/a	n/a	200
	28S	GTR+ Γ	n/a	n/a	300
	18S28S	GTR+ Γ	n/a	n/a	250
	1a18S28S	GTR+I+ Γ	n/a	n/a	250
	1a3 rd 18S28S	GTR+I+ Γ	n/a	n/a	250
<u>MrBayes (MLE)</u> McMC parameters: nchains=4; Temp=0.1; Swapfreq=1; Nswaps=1 4 gamma categories; estimated empirical base frequency EF1 α = 4 runs each to check for convergence 18S, 28S & concatenated = ≥ 2 runs each to check for convergence	1a	GTR+I+ Γ	*30	0.5	60,000
	1a3 rd	GTR+I+ Γ	*30	0.5	60,000
	18S	GTR+ Γ	10	0.5	5,000
	28S	GTR+ Γ	10	0.5	5,000
	18S28S	GTR+ Γ	10	0.5	5,000
	1a18S28S	GTR+I+ Γ	4	0.5	2,000
	1a3 rd 18S28S	GTR+I+ Γ	4	0.5	2,000

Table 3.4. Summary of recovered clades phylogenetic analyses based on previous results from Triemer et al. (2006). A ++ denotes a strongly supported clade (both analyses-types recover it), + denotes a supported clade (one analysis type recovers it) and a blank spot denotes it was not ever recovered or it was not possible to recover it due to reduced taxa studied this analysis compared to the Triemer et al. (2006) analysis major clades. Operational taxonomic units (OTUs) in the clades are either combinations of genera, a genus, or the species that represent a clade nested within a genus that are represented by two or more species in this study. When three or more species are represented in a clades by species sampled in this study, the organisms are described in Newick format to illustrate recovered relationships between them, although only the organisms together in a monophyletic clade is what was required for a +. A dash (-) represents that the OTUs in the clade were not sampled in this study. Analysis types were RAxML (ML; “best tree” used with rapid bootstrap support) and MrBayes (MLE; consensus tree). Genes are EF1 α =1a or EF1 α with third codon positions removed=1a3rd, small subunit ribosomal DNA=18S, large subunit ribosomal DNA=28S and the concatenated alignments utilizing EF1 α , 18S and 28S are either 18S28S, 1a18S28S, or 1a3rd18S28S.

Triemer et al. (2006) Clade	OTUs in Clade	Alignment						
		1a	1a3rd	18S	28S	18S28S	1a18S28S	1a3rd18S28S
Euglenea		+	++	++	++	++	++	++
root	Eutreptiales			++	++	++	++	++
A	<i>Discoplastis</i> (<i>Discoplastis adunca</i>)							
B	all other Euglenea (not A or root)			++				
C	<i>Phacus</i> + <i>Lepocinclis</i>			++			++	+
D	<i>Lepocinclis</i>			++		++	++	++
*D1	((<i>L. acus</i> var. <i>major</i> , <i>L. tripteris</i>), <i>L. spirogyroides</i>)							
*D2	<i>L. steinii</i> + <i>L. ovum</i> var. <i>ovum</i>			++		++	++	
E (=E1)	<i>Phacus</i> (<i>P. orbicularis</i> + <i>P. similis</i>)			++			++	+
*E2	-							
F	all other Euglenea (not A, B, C, D, E, or root)			++				
G	<i>Euglena</i>							
H	“H” subclade of <i>Euglena</i>			++	++	++	++	++
*H1	<i>E. stellata</i> + <i>E. geniculata</i>							
*H2	<i>E. granulata</i> + <i>E. viridis</i>							
*	((H1, H2), <i>E. deses</i>)							
*H3	(((<i>E. hiemalis</i> , <i>E. longa</i>), <i>E. gracilis</i>), <i>E. clara</i>)	++						
I	“I” subclade of <i>Euglena</i> (<i>E. exilis</i>)							
J	(K,L)					++		++
K	“mucilaginous clade” = ((<i>Strombomonas</i> , <i>Trachelomonas</i>), <i>Colacium</i>)					++	+	++
*	<i>Colacium</i> ((<i>C. mucronatum</i> , <i>C. elongatum</i>), <i>C.</i>			++	++	++	+	++

	<i>vesiculosum</i>) <u>Note:</u> <i>Colacium</i> sp. not in Triemer et al. (2006) study							
L	<i>Cryptoglana</i> + <i>Monomorphina</i>				++	++	++	++
*	<i>Cryptoglana</i>	++		++	++	++	++	++
*	<i>Monomorphina</i> ((<i>Monomorphina</i> sp., <i>M. pyrum</i>), <i>M. aenigmatica</i>)			++	++	++	++	++
M	“loricate clade” = <i>Strombomonas</i> + <i>Trachelomonas</i>							
*	<i>Strombomonas</i> ((<i>S. acuminata</i> , <i>S. triquetra</i>)(<i>S. borysteniensis</i> , <i>S. ovalis</i>)) <u>Note:</u> <i>S. napiformis</i> not in Triemer et al. (2006) study			++	++	++	++	++
N	<i>Trachelomonas</i>					++	++	++
*N1	<i>T. echinata</i>							
*N2	<i>T. similis</i> + <i>T. volvocina</i>	++		++	++	++	++	++
*N3	((<i>T. planctonica</i> , <i>T. lefevrei</i>), <i>T. conradi</i>)			++	++	++	++	++
Pluses total scores; % recovery (out of 26 possible)		4; 15%	1; 4%	15; 58%	10; 38%	15; 58%	16; 61%	16; 61%

Figure 3.1. The 18S interleaved alignment. Kinetoplastids are listed first and then euglenids, with an identifying number for each OTU to the left of the species name. The column positions, from 1-1324, are given in ten base increments at the top of each page. The number of residues from the original sequence is given in brackets to the right of each line corresponding to its respective OTU.

	10	20	30	40	50	60	70	80	90	100	
01_Rhynchomonas_nasuta	TgCcAgTaGt	CaTaTgCtTg	TtTcAaGgAc	TaAgCcAtGc	AtGcCtCaGa	AtAaCtGcCa	GgAgTcTgGg	CaTgGcTcAt	TaCaTcAgAc	GcAaTcTcGc	[100]
02_Blastocrithidia_culicis	-----	CATATGCTTG	TTTCAAGGAC	TAAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[90]
03_Herpetomonas_muscarum	-----	CATATGCTTG	TTTCAAGGAC	TTAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[90]
04_Leishmania_braziliensis	-----TAGT	CATATGCTTG	TTTCAAGGAC	TTAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[94]
05_Leptomonas_bifurcata	-----	CATATGCTTG	TTTCAAGGAC	TTAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[90]
06_Leishmania_infantum	TGCCAGTAGT	CATATGCTTG	TTTCAAGGAC	TTAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[100]
07_Leishmania_major	-----TAGT	CATATGCTTG	TTTCAAGGAC	TTAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[94]
08_Leishmania_tarentolae	TGCCAGTAGT	CATATGCTTG	TTTCAAGGAC	TTAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[100]
09_Trypanosoma_cruzi	-----GT	CGTATGCTTG	TCTCAAGGAC	TTAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TGCATCAGAC	GTAATCTCGC	[92]
10_Trypanosoma_brucei	TGCCAGTAGT	CATATGCTTG	TTTCAAGGAC	TTAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[100]
11_Bodo_saltans	TGCCAGTAGT	CATATGCTTG	TTTCAAGGAC	TAAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[100]
12_Dimastigella_trypaniformis	-----GT	CATATGCTTG	TTTCAAGGAC	TAAGCCATGC	ATGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[92]
13_Rhynchobodo_sp	-----T	CATATGCTTG	TTTCAAGGAC	TAAGCCATGC	ACGCCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGAC	GTAATCTCGC	[91]
14_Colacium_elongatum	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCCAGT	[100]
15_Ploeotia_oblonga	TGCAGTAGCC	-GTATGCTTG	TTTCAAGGAC	TAAGCCATGC	AAACCTCAGA	ATCACTGCCA	GGAATCTGCG	CATGGCTCAT	TACATCAGCA	GTCATCCAGT	[97]
16_Cryptoglena_pigra	-----	-----	-----	-----ATGC	ATGTCTCAGA	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTTGT	[64]
17_Cryptoglena_skujai	-----	-----	-----	-----ATGC	ATGTCTCAGA	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTTGT	[64]
18_Colacium_sp	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCCAGT	[100]
19_Colacium_vesiculosum	-----	-----	-----	-----	-----	-----	-----	-----T	TACATCAGCA	GTCATCCAGT	[21]
20_Euglena_clara	TCCCAACAG-	GGTGGTCC	TTTGTATGCA	GGTGTGTGCG	CTGATTGCGC	CCAGTGCAGC	CCGCGCTCGG	ATTTGATCGG	AAGAGCAGCG	TT--CCAGT	[97]
21_Discoplastis_adunca	-----CAGC	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGA	GTACATGGAG	TGGATCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTTGT	[94]
22_Euglena_deses	-----T--AGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	AACTCTCAGC	GCACATGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGC	[94]
23_Euglena_geniculata	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCCAGT	[100]
24_Euglena_granulata	-----CAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[94]
25_Euglena_hiemalis	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ACGTCTCAGC	GCAAAACGGAG	TGGATCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[100]
26_Eutreptia_pertyi	-----	-----	-----TCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GTACACGGAG	TGGATCTGTG	AATGGCTCCT	TACATCAGCA	ACAATCGTAC	[78]
27_Euglena_stellata	-----	-----C	GATCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTTGT	[81]
28_Euglena_longa	-----	-----	-----	-----ATGC	ACGTCTCAGC	GCAAAACGGAG	TGGATCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[64]
29_Euglena_viridis	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GTACACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTTGT	[100]
30_Euglena_exilis	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATATCTCAGC	GTACACGGAG	TGGATCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGC	[100]
31_Euglena_gracilis	-----	-----	-----	-----ATGC	ACGTCTCAGC	GCAAAACGGAG	TGGATCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[64]
32_Lepocinlis_steinii	-----	-----	-----	-----ATGC	ATGTGTGAGA	GTACACG-AG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GGCATCTTGT	[63]
33_Lepocinlis_acus_var_major	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GTACACG-AG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[99]
34_Lepocinlis_tripteris	-----	-----	-----	-----	-----	-----	-----	-----T	TACATCAGCA	GTCATCTCGT	[21]
35_Lepocinlis_spirogyroides	-----	-----	-----	-----	-----	-----	-----	-----T	TACATCAGCA	GTCATCTAGT	[21]
36_Lepocinlis_ovum	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GTACACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GGCATCTCGT	[100]
37_Monomorphina_aenigmatica	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ACGTCTCAGC	GTACACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATGTCGT	[100]
38_Colacium_mucronatum	-----	-----	-----CAAGGAC	TAAGCCATGC	ATGTCTCTGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCCAGT	[77]
39_Monomorphina_sp	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGT	GTACACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTTGT	[100]
40_Peranema_trichophorum	TGCCAGTAGT	CATATGCTTG	TTTCAAGGAC	TAAGCCATGC	AAGTGTAAAG	GTAAATGCCA	TAGGTCTGCG	AATGGCTCAT	TACAACAGCA	ACAATCCATT	[100]
41_Phacus_orbicularis	-----	-----	-----	-----ATGC	AAGTCTCAGC	GTACACG-AG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGC	[63]
42_Phacus_similis	-----	-----	-----	-----ATGC	ATGTCTCAGT	GTACACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCAT	[64]
43_Monomorphina_pyrum	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGT	GTACACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTTGT	[100]
44_Strombomonas_acuminata	-----	-----	-----	-----	-----	-----	-----	-----T	TACATCAGCA	GTCATCTAGT	[21]
45_Strombomonas_napiformis	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTAGT	[100]
46_Strombomonas_borysteniensis	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCCAGT	[100]
47_Strombomonas_ovalis	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCCAGT	[100]
48_Trachelomonas_echinata	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTTGT	[100]
49_Strombomonas_triquetra	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTAGT	[100]
50_Eutreptiella_pomquetensis	-----	-----	-----	-----ATGC	AAGCTCTCAG	GTACATGGAG	TAGATCTGTG	AATGGCTCAT	TACATCAGCA	ATAATCTCCT	[64]
51_Trachelomonas_lefevrei	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCACAATGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[100]
52_Trachelomonas_planctonica	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCACATGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[100]
53_Trachelomonas_similis	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[100]
54_Trachelomonas_volvocina	-----	-----	-----	-----ATGC	ATGTCTCAGC	GCAAAACGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[64]
55_Trachelomonas_conradi	TGCCAGTAGT	CATATGCTTG	GTTCAAGGAC	TAAGCCATGC	ATGTCTCAGC	GCACATGGAG	TGGGTCTGTG	AATGGCTCCT	TACATCAGCA	GTCATCTCGT	[100]

	110	120	130	140	150	160	170	180	190	200	
01_Rhynchomonas_nasuta	AaAaGcGaAa	CgCcAaGcTa	AtAcAtGaAc	TaAgCtCgGa	AaGaAcGcGc	TgAaTcAaTt	-tTtTgA-Cg	AaCaAcTgCc	CtAtCaGcCg	AtGgCcGtGt	[198]
02_Blastocrithidia_culicis	AAGAGCGAAA	CGCCTAGCTA	ATACATGAAC	CAAGCGCTGA	CCGATGCGGT	TGACTCAATT	TTTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[189]
03_Herpetomonas_muscarum	AAAAGCGAAA	AG-CAAGCTA	ATACATGAAC	CAAGCGTTGA	CATAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[187]
04_Leishmania_braziliensis	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[192]
05_Leptomonas_bifurcata	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[186]
06_Leishmania_infantum	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AAC-ACTGCC	GTATCAATCG	ATGCC-CTGT	[196]
07_Leishmania_major	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[192]
08_Leishmania_tarentolae	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[198]
09_Trypanosoma_cruzi	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[190]
10_Trypanosoma_brucei	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[199]
11_Bodo_saltans	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[198]
12_Dimastigella_trypaniformis	AAAAGCGAAA	CGCCAAGCTA	ATACATGAAC	CAAGTGTGTA	CAGAATGCGT	TGACTCAATT	-TTTTGA-CG	AACAACCTGCC	CTATCAGCTG	ATGGCCGTGT	[191]
13_Rhynchobodo_sp	AAAAGCGAAA	CG-CAAGCTA	ATACATGAAC	TTAATGTCGA	AATAACG-GT	TGAATCAATT	TTTTCGA-CG	AACAACCTGCC	CTAT- - - - -	- - - - -CGTGT	[177]
14_Colacium_elongatum	GGAACAAAAC	CTGGTGGCTA	ATACATGAAC	GA-CCGCCGT	GA-CCGCCGT	TGACTCAGGC	TGATCAGAGG	AACCTCTGAC	CTATCAGCTG	ATTGTGGTGT	[198]
15_Ploeotia_oblonga	GTTCACAAAC	TGTATCGCTA	ATACGTGTAT	GAGCTGGAGC	AGTTAACGGG	TTACAATGGC	-AACTGGACG	AATTTCTGCC	TTATCAGCTG	ATGACTGTGC	[196]
16_Cryptoglena_pigra	GATACAAAAC	CTGGTGGCTA	ATACACGTTG	GACCTGCCGC	TTTGGC-AGT	TGACTCAGGC	TGATCACAAG	AACCTCTGCC	CTATCAGCTG	ACTGTGGTGT	[163]
17_Cryptoglena_skujai	GATACAAAAC	CTGGTGGCTA	ATACACGTTG	GACCTGCCGC	TCTGGC-AGC	TGACTCAGGC	TGATCACAAG	AACCTCTGCC	CTATCAGCTG	ACTGTGGTGT	[163]
18_Colacium_sp	GGTACAAAAC	CTGGTGGCTA	ATACACGGAA	GA-CCGCCGT	TTGAGC-GGC	TGAATCAGGC	TGACCACAGG	AACCTCTGAC	CTATCAGCTG	ATTGTGGTGT	[198]
19_Colacium_vesiculosum	GGTACAAAAC	CTGGTGGCTA	ATACACGGAA	GA-CTGCCGT	TTGAGC-GGC	TGAATCAGGC	TGACCACAGG	AACCTCTGAC	CTATCAGCTG	ATTGTGGTGT	[119]
20_Euglena_clara	GGTTAGCAGT	CGCCCGCGCA	CCACCCACAC	ACCCGCCGTG	AAGTGCCTTC	GGTTTCGGGA	TGATCAGCGC	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[197]
21_Discoploplastis_adunca	GGTGCAAAAAC	CTAGTGGCTA	ATACACGGTC	GAACATATCGA	ATGAGT-GGG	TGACTCAGGC	TGACCACATG	AACCTCTGAC	CTATCAGCTG	ATTGTGGTGT	[193]
22_Euglena_deses	GATACAAAAC	CTTGTGGCTA	ATACACGTTT	GCTCGGCCGG	TATAGCGTGT	TGACCCAGGC	TGATCGCACG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[194]
23_Euglena_geniculata	GACAGAAAAC	CTTGTGGCTA	ATACACGTTT	GCCCGGCTGG	TCGAGGCGAGT	TGATTCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ATTGTGGTGT	[200]
24_Euglena_granulata	GATACAAAAC	CTGGTGGCTA	ATACACGCGA	-GCCGGCCGG	GCGAGT-GTC	CGTGTCGGGG	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[192]
25_Euglena_hiemalis	GATACAAAAC	CTTGTGGCTA	ATACACGCGT	GAACAGCCGA	TACAGCGGGT	TGACTCAGGC	TGATCGCAAG	AGCTTCTGAC	CTATCAGCTG	ACTGTGGTGT	[200]
26_Eutreptia_pertyi	GGAGACCAAC	TGGCTGGCTA	ATACACGGAA	GTTCTCCTGT	GTTTGT- -GA	TGATTC-AGA	TCCTGGAGCG	AACCTCTGAC	CTATCAGCTG	ATTGTGGTGT	[175]
27_Euglena_stellata	GATCCAAAAC	CTTGAGGCTA	ATACACGAAT	GCACGGCCGG	TATAGT-GGT	TGACTCAGGC	TGATCGCTCG	AACCTCTGAC	TTATCAGCTG	ATTGTGGTGT	[180]
28_Euglena_longa	GATACAAAAC	CTTGTGGCTA	ATACACGCTC	GAACAGCCGA	TACAGCGGGT	TGACTCAGGC	TGATCGCAAG	AGCTTCTGAC	CTATCAGCTG	ACTGTGGTGT	[164]
29_Euglena_viridis	GATACATAAC	CTTGTGGCTA	ATACACGAAT	GCCCGGCCGA	TATAGC-GGT	TGATTCAGGC	TGATCACAATG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[199]
30_Euglena_exilis	GAACAAAAC	CTGGTGGCTA	ATACACGTTT	GC-CTGCCGT	TTGAGC-GGT	TGACTCAGGC	TGATCGCACG	AACCTCTGAC	CTATCAGCTG	ACTGTGGCGC	[198]
31_Euglena_gracilis	GATACAAAAC	CTTGTGGCTA	ATACACGCGT	GAACAGCCGA	TACAGCGGGT	TGACTCAGGC	TGATCGCAAG	AGCTTCTGAC	CTATCAGCTG	ACTGTGGTGT	[164]
32_Lepocinclis_steinii	GATACAAAAC	CTAGTGGCTA	ATACACGGTC	GTACTGCTTT	CTGAGC-GGT	TGACTC-AGG	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[161]
33_Lepocinclis_acus_var_major	GATACAAAAC	CTAGTGGCTA	ATACACGGTT	GCCCTGCTGC	CTTAGC-GGT	TGACTCAGGC	TGATCACAAG	AGCTTCTGAC	CTATCAGCTG	ATGGTGGTGT	[198]
34_Lepocinclis_tripteris	GATACAAAAC	CTAGTGGCTA	ATACACGTTT	GCCCTGCCGA	TTTAGC-GGT	TGACTCAGGC	TGATCACAAG	AACCTCTGAT	CTATCAGCTG	ATGGTGGTGT	[120]
35_Lepocinclis_spirogyroides	GATACAAAAC	CTAGTGGCTA	ATACACGCAA	GCCCGGCCGC	TTGAGC-GGT	TGACTCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ATGGTGGTGT	[120]
36_Lepocinclis_ovum	GATACAAAAC	CTAGTGGCTA	ATACACGGTT	GCACTGCTGT	CTGAGT-GGC	TGACTCAGGC	TGATCGCACG	AGCTTCTGAC	CTATCAGCTG	ATGGTGGTGT	[199]
37_Monomorphina_aenigmatica	GATACAAAAC	CTGGTGGCTA	ATACACGGAT	GGACTGCCGC	TTGAGC-GGT	TGACTCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[199]
38_Colacium_mucronatum	GGTACAAAAC	CTGGTGGCTA	ATACACGGAA	GA-CTGCCGT	TTGAGC-GGC	TGAATCAGGC	TGACCACAGG	AACCTCTGAC	CTATNAGCTG	ATTGTGGTGT	[175]
39_Monomorphina_sp	GATACAAAAC	CTGGTGGCTA	ATACACGGAA	GCACCCCTGT	CTGATGCGGT	TGACTCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[200]
40_Peranema_trichophorum	GACTCTCAAC	CGGCTAGTTA	ATACGTGTAC	GAAACT- -GA	CCTGGT-GGT	TGAGTC-AGT	AAGTCATAGG	AACCTCTGCC	CTATCAGCTG	ATGGTGAGGT	[196]
41_Phacus_orbicularis	GATACAAAAC	CTAGTGGCTA	ATACACGGTC	GACCTGCTGC	TCGAGC-GGC	TGACTCAGGC	TGATCGCACG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[162]
42_Phacus_similis	GATACAAAAC	CTAGTGGCTA	ATACACGGTC	GAGCTGCCGC	TGAAGC-GGC	TGACTCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[163]
43_Monomorphina_pyrum	GATACAAAAC	CTGGTGGCTA	ATACACGGAA	GCACCCCTGT	CTGAGC-GGT	TGACTCAGGC	TGATCGCAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[199]
44_Strombomonas_acuminata	GATGCAAAAAC	CTAGTGGCTA	ATACACGAAC	GGACTGCCGC	CTTTGCGGGC	TGACTCAAGC	TGATCACTCG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[121]
45_Strombomonas_napiformis	GATGCAAAAAC	CTAGTGGCTA	ATACACGAAA	GCACTGCCGC	CTTTGCGGGC	TGACTCAAGC	TGATCACTCG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[200]
46_Strombomonas_borysteniensis	GATGCAAAAAC	CTAGTGGCTA	ATACACGGAA	GCACTGCCGC	CTTTGCGGGC	TGACTCAGGC	TGATCACTCG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[200]
47_Strombomonas_ovalis	GATGCAAAAAC	CTAGTGGCTA	ATACACGGAA	GCACTGCCGC	CTTTGCGGGC	TGACTCAGGC	TGATCACTCG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[200]
48_Trachelomonas_echinata	GATAAATAAC	CTTGTGGCTA	ATACACGCAA	GCACTGCCCTC	CTGAGCACGG	TGAATCAGGC	TGATCACAATG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[200]
49_Strombomonas_triquetra	GATGCAAAAAC	CTAGTGGCTA	ATACACGGAT	GAACCTGCCGA	CTTTGCGGGC	TGACTCAGGC	TGATCACTCG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[200]
50_Eutreptiella_pomquetensis	GGTGCTTAAC	TAGGTGGTTA	ATACATGGAA	GAACCTGCCGA	TACGGC-GGG	TGATTCAGGC	TGACCACAGG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[163]
51_Trachelomonas_lefevrei	GATACAAAAC	CTGGTGGCTA	ATACACGAAA	GCACTGCCGC	ATGAGC-GGC	TGAATCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[199]
52_Trachelomonas_planctonica	GATACAAAAC	CTGGTGGCTA	ATACACGAAA	GCACTGCCGC	ATGAGC-GGC	TGAATCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[199]
53_Trachelomonas_similis	GATAGAAAAC	CTTGTGGCTA	ATACACGGAA	GCACTCTCAG	CTAGGC-GGT	TGAATCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[199]
54_Trachelomonas_volvocina	GATAAAAAAC	CTTGTGGCTA	ATACACGGAA	GCACGGATGA	GAGAGTGC	TGAATCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ACTGTGGTGT	[164]
55_Trachelomonas_conradi	GATACAAAAC	CTGGTGGCTA	ATACACGAAT	GGACTGCCGC	ATGAGTGC	TGACTCAGGC	TGATCACAAG	AACCTCTGAC	CTATCAGCTG	ATTGTGGTGT	[200]

	210	220	230	240	250	260	270	280	290	300	
01_Rhynchomonas_nasuta	AgTgGaCtGc	CaTgGcGtTg	AcGgG-AgCg	GgGgAtTaGg	GtTcGaTtCc	GgAgAgGgAg	CcTgAgAaAt	GgCtAcCaCt	TcTaCgGaGg	GcAgCaGgCg	[297]
02_Blastocrithidia_culicis	AGTGGACGGC	CATGGCGTTG	ACGGG-AGCG	GGGAAATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[288]
03_Herpetomonas_muscarum	AGTGGACTGC	CATGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[286]
04_Leishmania_braziliensis	AGTGGACTGC	CATGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[291]
05_Leptomonas_bifurcata	AGTGGACTGC	CAGGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[285]
06_Leishmania_infantum	AATGGACTGC	CAT-GCTTTT	ACCGA--ACG	GGGGAATAAG	G-TCGATTTC	GGAAAGGAGG	CTGG-AAAAT	TGCTT-CACC	TCTACCGAAG	GGCACAAGCG	[290]
07_Leishmania_major	AGTGGACTGC	CATGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[291]
08_Leishmania_tarentolae	AGTGGACTGC	CATGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[297]
09_Trypanosoma_cruzi	AGTGGACTGC	CATGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	CCTACGGAGG	GCAGCAGGCG	[289]
10_Trypanosoma_brucei	AGTGGACTGC	CATGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[298]
11_Bodo_saltans	AGTGGACTGC	CATGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	AGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[297]
12_Dimastigella_trypaniformis	AGTGGACTGC	CATGGCGTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	GGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[290]
13_Rhynchobodo_sp	AGTGGACTGC	CATGGCTTTG	ACGGG-AGCG	GGGGATTAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	GGCTACCACT	TCTACGGAGG	GCAGCAGGCG	[276]
14_Colacium_elongatum	AGTGGACCAC	AATGGCATTG	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[298]
15_Ploeotia_oblonga	AGTGAACAGT	CATGGTATTA	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAC	GGCGACCCAGA	TCTAAGGTATT	GCAGCAGGCA	[295]
16_Cryptoglena_pigra	AGTGGACCAC	AGTGGCAATA	ACGGGTAACG	GAGAAATCTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[263]
17_Cryptoglena_skujai	AGTGGACCAC	AGTGGCAGTA	ACGGGTAACG	GAGAAATCTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[263]
18_Colacium_sp	AGTGGACCAC	AATGGCGTCA	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[298]
19_Colacium_vesiculosum	AGTGGACCAC	GATGGCATTG	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[219]
20_Euglena_clara	AGCGGACACA	AGTGGCAGTT	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[297]
21_Discoplastis_adunca	AGTGGACCAC	AATGGCGTCA	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAT	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[293]
22_Euglena_deses	CGTGGACCAC	AGTGGCCTTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[294]
23_Euglena_geniculata	AGTGGACCAC	AATGGCTTTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[300]
24_Euglena_granulata	AGTGGACCAC	AGTGGCGGTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[292]
25_Euglena_hiemalis	ATTGGACCAC	AGTGGCCTTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	TCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCA	[300]
26_Eutreptia_pertyi	CGTGGACTAC	AATGGCTTCC	ACGGGTAACG	GAGAAATGAGG	GTTTGTATACC	GGAGAGGGGAG	CCTGAGAGAT	GGCTACCACT	GCTAAGGCGG	GCAGCAGGCG	[275]
27_Euglena_stellata	AGTGGACCAC	AATGGCTTCT	ACGAGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[280]
28_Euglena_longa	ATTGGACCAC	AGTGGCCTTG	ACGGGTAACG	AAGAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCA	[264]
29_Euglena_viridis	AGTGGACCAC	AGTGGCTATG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAT	GGCTACCACT	ACTAAGGTGA	GCAGCAGGCG	[299]
30_Euglena_exilis	AAGGGCCAC	AGTGGCGGCA	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[298]
31_Euglena_gracilis	ATTGGACCAC	AGTGGCCTTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCA	[264]
32_Lepocinclis_steinii	CGAGGCC-AC	CATGGCAGTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACTAAGGTGG	GCAGCAGGCA	[260]
33_Lepocinclis_acus_var_major	ATTGGACCAC	CATGGCAGTG	ACGGGTAACG	GAGAAATCAGG	GTTTGTATACC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	TCTAAGGAGG	GCAGCAGGCA	[298]
34_Lepocinclis_tripteris	AGTGGACCAC	CATGGCAGTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAT	GGCTACCACT	TCTAAGGAGG	GCAGCAGGCA	[220]
35_Lepocinclis_spirogyroides	AGTGGACCAC	CATGGCGGTG	ACGGGTAACG	GAGAAATCAGG	GTTTGTATACC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	TCTAAGGAGG	GCAGCAGGCA	[220]
36_Lepocinclis_ovum	AGAGCCC-AC	CATGGCGGTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	TCTAAGGAGG	GCAGCAGGCA	[298]
37_Monomorphina_aenigmatica	AGTGGACCAC	AGTGGCGGTG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCA	[299]
38_Colacium_mucronatum	AGTGGACCAC	AATGGCATTG	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTTAGAGAC	GGCTAC-CCT	ACCAAGGTGG	GCAGCAGGCG	[274]
39_Monomorphina_sp	AGTGGACCAC	AGTGGCAGTA	ACGGGTAACG	GAGAAATCTGG	GTTTGTATACC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[300]
40_Peranema_trichophorum	AGTGGCTCAC	CATGGCATTG	ACGGGTAACG	GAGAAATAGG	GTTTGTATACC	GGAGAGGGGAG	CCTGAGAAAT	GGCTACCACT	GCTATGGCAG	GCAGCAGGCG	[296]
41_Phacus_orbicularis	AGTGGACCAC	AGTGGCTATG	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	GCTAAGGCGG	GCAGCAGGCA	[262]
42_Phacus_similis	AGTGGACCAC	AGTGGCAGTT	ACGGGTAACG	GAGAAATCAGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[263]
43_Monomorphina_pyrum	AGTGGACCAC	AGTGGCCGTA	ACGGGTAACG	GAGAAATCTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTNAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[299]
44_Strombomonas_acuminata	AGTGGACCAC	AGTGGCCATG	ACGGGTAACG	GAGAAATTTGG	GTTTCGACTCC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[221]
45_Strombomonas_napiformis	AGTGGACCAC	AGTGGCCGTT	ACGGGTAACG	GAGAAATTTGG	GTTTCGACTCC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[300]
46_Strombomonas_borysteniensis	AGTGGACCAC	AGTGGCAATC	ACGGGTAACG	GAGAAATTTGG	GTTTCGACTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[300]
47_Strombomonas_ovalis	AGTGGACCAC	AGTGGCAGTG	ACGGGTAACG	GAGAAATTTGG	GTTTCGACTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[300]
48_Trachelomonas_achinata	AGTGGACCAC	AGTGGCCCTG	ACGGGTAACG	GAGAAATCAGG	GTTTGTATACC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[300]
49_Strombomonas_triquetra	AGTGGACCAC	AGTGGCCATG	ACGGGTAACG	GAGAAATTTGG	GTTTCGACTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[300]
50_Eutreptiella_pomquetensis	AGTGGACCAC	CGTGGCTATA	ACGGGTAACG	GAGAAATCAGG	GTTTGTATACC	GGAGAGGGGAG	CCTGATAAAC	GGCTACCACT	GCTAAGGTGG	GCAGCAGGCA	[263]
51_Trachelomonas_lefevrei	AGTGGACCAC	AGTGGCAGTG	ACGGGTAACG	GAGAAATTTGG	GTTTGTATACC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[299]
52_Trachelomonas_planctonica	AGTGGACCAC	AGTGGCAGTG	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAAAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[299]
53_Trachelomonas_similis	AGCGGACCAC	GTTGGCTATG	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	GCTAAGGTGG	GCAGCAGGCG	[299]
54_Trachelomonas_volvocina	AGCGGACCAC	AGTGGCTTTG	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	GCTAAGGTGG	GCAGCAGGCG	[264]
55_Trachelomonas_conradi	AGTGGACCAC	AATGGCAGTG	ACGGGTAACG	GAGAAATTTGG	GTTTCGATTCC	GGAGAGGGGAG	CCTGAGAGAC	GGCTACCACT	ACCAAGGTGG	GCAGCAGGCG	[300]

	310	320	330	340	350	360	370	380	390	400	
01_Rhynchomonas_nasuta	CgCaAaTtGc	CcAaTgCaTg	AgGcAgCgAa	AaGaAaTaGg	TgCcCcGgGg	CtTtAaTgGg	GgCaTcGaGt	AaCaAtTgGa	GgAcAaGtCt	GgTgCcAgCa	[397]
02_Blastocrithidia_culicis	CGCAAATTGC	CCAATGTATG	AGGCAGCGAA	AAGAAATAGG	GTTGTCGGCA	-TTTAATGGG	GGTATCGAGT	AGCCATTGGA	GGACAAGTCT	GGTGCCAGCA	[387]
03_Herpetomonas_muscarum	CGCAAATTGC	CCAATGTATG	AGGCAGCGAA	AAGAAATAGG	GTTGTCGTGA	-TTTAATGGG	GGTATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[385]
04_Leishmania_braziliensis	CGCAAATTGC	CCAATGTATG	AGGCAGCGAA	AAGAAATAGA	GTTGTCGTGA	-TTTAATGGG	GGTATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[390]
05_Leptomonas_bifurcata	CGCAAATTGC	CCAATGCATG	AGGCAGCGAA	AAGAAATAGA	GTTGTCGTGA	-TTTAATGGG	GGTATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[384]
06_Leishmania_infantum	CCCAAATGGT	CC-ATGTATG	AGGCCGCGAA	-AGAAAAA	G-TGTCGTCA	-TTT-ATGGG	GGCATCCGGG	-ACATTTGTA	GGAC--ATCT	TGTGCCAGCA	[382]
07_Leishmania_major	CGCAAATTGC	CCAATGTATG	AGGCAGCGAA	AAGAAATAGA	GTTGTCGTGA	-TTTAATGGG	GGTATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[390]
08_Leishmania_tarentolae	CGCAAATTGC	CCAATGTATG	AGGCAGCGAA	AAGAAATAGA	GTTGTCGTGA	-TTTAATGGG	GGTATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[396]
09_Trypanosoma_brucei	CGCAAATTGC	CCAATGTATG	AGGCAGCGAA	AAGAAATAGA	GCCGACGTG	TTTTAATGGG	GGAATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[389]
10_Trypanosoma_brucei	CGCAAATTGC	CCAATGTATG	AGGCAGCGAA	AAGAAATAGA	GCCGACGTG	TTTTAATGGG	GGTATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[398]
11_Bodo_saltans	CGCAAATTGC	CCAATGTATG	AGGCAGCGAA	AAGAAATGGC	CTCC-CGGGA	GATAAATGGG	GGAATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[396]
12_Dimastigella_trypaniformis	CGCAAATTGC	CCAATGCATG	AGGCAGCGAA	AAGAAATAGG	TGCCCTAGGG	CTTTAATGAC	CGCATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[390]
13_Rhynchobodo_sp	CGTAAATTGC	TCAATGCATG	AGGCAGCGAA	AAGAAATAGG	TTTTCCGGAA	TTTAAATGGG	GGTATCGAGT	AACAATTGGA	GGACAAGTCT	GGTGCCAGCA	[376]
14_Colacium_elongatum	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACCGCAAC	AATCCTAGGA	CTTGAATGGA	TGTCATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[398]
15_Ploeotia_oblongata	CGCAAATTGC	CCAATGCTTG	AGGCAGCGAG	GAGAAATAAC	AACACCGGTG	ATTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[395]
16_Cryptoglana_pigra	CGCAAATTGC	CCAATCCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTCATGAGT	ACCTATTGGA	GGGCAAGTCT	GGTGCCAGCA	[363]
17_Cryptoglana_skujai	CGCAAATTGC	CCAATCCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTCATGAGT	ACCTATTGGA	GGGCAAGTCT	GGTGCCAGCA	[363]
18_Colacium_sp	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACCGCAAC	AATCCTAGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[398]
19_Colacium_vesiculosum	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACCGCAAC	AATCCTAGGA	CTTGAATGGA	TGTCATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[319]
20_Euglena_clara	CGCAAATTGC	CCAATCCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCG- GA	CTTGAATGGA	TGTGATGAGT	AACAAGTCTGGA	GGGCAAGTCT	GGTGCCAGCA	[396]
21_Discoplastis_adunca	CGCAAATTGC	CCAATGC- CG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTG-ATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[392]
22_Euglena_deses	CGCAAATTGC	CCAATGC- CG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTG-ATGGA	TGTGATGAGT	AACAAGTCTGGA	GGGCAAGTCT	GGTGCCAGCA	[392]
23_Euglena_geniculata	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[400]
24_Euglena_granulata	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTG-ATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[391]
25_Euglena_hiemalis	CGCAAATTGC	CCCATGCGTG	AGGCAGCGAG	GAACAGTAGC	AATCCTGGGA	CTTGAATGGA	TGTCATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[400]
26_Eutreptia_pertyi	CGCAAATTGC	CCTATACGTG	AGGCAGCAAC	GAGTCGTGGC	AGTTCTAGGA	CTGGAATGGA	TGCAATGAGC	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[375]
27_Euglena_stellata	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GATCCGCAAGC	AATCCTGGGA	CTTG-ATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[379]
28_Euglena_longa	CGCAAATTGC	CCCATGCGTG	AGGCAGCGAG	GAACAGTAGC	AATCCTGGGA	CTTGAATGGA	TGTCATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[364]
29_Euglena_viridis	CGCAAATTGC	CCAATGC- CG	AGGCAGCGAG	GAACGCAAGC	AATCCTAGGA	CTTG-ATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[397]
30_Euglena_exilis	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[398]
31_Euglena_gracilis	CGCAAATTGC	CCCATGCGTG	AGGCAGCGAG	GAACAGTAGC	AATCCTGGGA	CTTGAATGGA	TGTCATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[364]
32_Lepocinclis_steinii	CGCAAATTGC	CCCATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[360]
33_Lepocinclis_acus_var_major	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACCGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[398]
34_Lepocinclis_tripteris	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[320]
35_Lepocinclis_spirogyroides	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[320]
36_Lepocinclis_ovum	CGCAAATTGC	CCCATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[398]
37_Monomorphina_aenigmatica	CGCAAATTGC	CCAATG- -CG	AGGCAGCAAC	GAACCGCAAGC	AATCCCGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[397]
38_Colacium_mucronatum	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACCGCAAC	AATCCTAGGA	CTTGAATGGA	TGTCATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[374]
39_Monomorphina_sp	CGCAAATTGC	CCAATACGTG	AGGCAGCAAC	GAACCGCAAC	AATCTTGGGA	TCTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[400]
40_Peranema_trichophorum	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	ATTTCAGAG	TTTGCATGGA	TATAATGAGT	AACAAGTCTGGA	GGGCAAGTCT	GGTGCCAGCA	[396]
41_Phacus_orbicularis	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCCCT- GA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[361]
42_Phacus_similis	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCTCA- GA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[362]
43_Monomorphina_pyrum	CGCAAATTGC	CCAATACGTG	AGGCAGCAAC	GAACCGCAAC	AATCTTGGGA	TCTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[399]
44_Strombomonas_acuminata	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTG-ATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[320]
45_Strombomonas_napiformis	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTG-ATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[399]
46_Strombomonas_borysteniensis	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[400]
47_Strombomonas_ovalis	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[399]
48_Trachelomonas_lefevrei	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[399]
49_Strombomonas_triquetra	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[400]
50_Eutreptiella_pomquetensis	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[363]
51_Trachelomonas_similis	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[399]
52_Trachelomonas_planctonica	CGCAAATTGC	CCAATACGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[397]
53_Trachelomonas_similis	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[399]
54_Trachelomonas_volvocina	CGCAAATTGC	CCAATGC- CG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[363]
55_Trachelomonas_conradi	CGCAAATTGC	CCAATGCGTG	AGGCAGCGAG	GAACGCAAGC	AATCTTGGGA	CTTGAATGGA	TGTGATGAGT	ATCAACTGGA	GGGCAAGTCT	GGTGCCAGCA	[400]

	410	420	430	440	450	460	470	480	490	500	
01_Rhynchomonas_nasuta	CcCgCgGtAa	TtCcAgCtCc	AaGaGcGtAt	AtTaAcGcTg	TtGcTgAaGg	GtTcGtAgTt	GaAtGgAgTc	TgTtCaTtAc	TgTgAcTaAa	GaAgTgCgAc	[497]
02_Blastocrithidia_culicis	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAACGGGCCG	TGATTT - TAC	TGTGACTAAA	GAAGTGCAC	[486]
03_Herpetomonas_muscarum	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAATGGGCCG	TGATTT - TAC	TGTGACTAAA	GAAGTGTGAC	[484]
04_Leishmania_braziliensis	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAACGGGCCG	TGATTTTAC	TGTGACTAAA	GAAGCGTGAC	[490]
05_Leptomonas_bifurcata	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAACGGGCCG	TGATTTTAC	TGTGACTAAA	GAAGTGTGAC	[484]
06_Leishmania_infantum	CCACC - GTAA	- TCCTGCTTC	AACCAAGG - GT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAACGGGCCG	TGATTTTAC	TGTGACTAAA	GAAGCGTGAC	[479]
07_Leishmania_major	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAACGGGCCG	TGATTTTAC	TGTGACTAAA	GAAGCGTGAC	[490]
08_Leishmania_tarentolae	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAACGGGCCG	TGATTTTAC	TGTGACTAAA	GAAGCGTGAC	[496]
09_Trypanosoma_cruzi	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAATGGGCCG	TGATTTTAC	TGTGACTAAA	AAAGTGTGAC	[489]
10_Trypanosoma_brucei	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAATGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAACGGGCCA	TGAATTTTAC	TGTGACCAAA	AAAGTGCAC	[498]
11_Bodo_saltans	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAACGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAACGGGCCG	TGATTT - TAC	TGTGACTAAA	GAAGCGCGAC	[495]
12_Dimastigella_trypaniformis	CCCGCGGTAA	TTCCAGCTCC	AAGAGCGTAT	ATTAACGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAATGGGTTT	TGTTTCATTAC	TGTGACTAAA	AAAGTGCAC	[490]
13_Rhynchobodo_sp	CCCGCGGTAA	TTCCAGCTCC	AAAAGCGTAT	ATTAACGCTG	TTGCTGAAGG	GTTTCGTAGTT	GAATGGGTTG	TGATATTAC	TGTGACAAAA	GAAGGGTGAC	[476]
14_Colacium_elongatum	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	GCTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCGGGGCC	TTGCTGTAC	CGTGGAAAA	TCAGTGTGCT	[498]
15_Ploeotia_oblonga	GTTAACGTTG	TTCCAAACAA	TGTGACACTA	CTGCTGAAAC	CGCGGAAAC	AAACAAAAAT	TGCCAGATGT	GTGGTGTTAC	CGTGAACAAA	CAAGCGTGTT	[495]
16_Cryptoglana_pigra	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCTGGGCAC	TCGTTGTAC	TGTGAAAA	TCAGTGTGCC	[463]
17_Cryptoglana_skujai	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGCAC	TCGTTGTAC	TGTGAAAA	TCAGTGTGCC	[463]
18_Colacium_sp	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	GCTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGGCC	TTGCTGTAC	CGTGGAAAA	TCAGTGTGCT	[498]
19_Colacium_vesiculosum	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	GCTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCGGGGCC	TTGATGTAC	CGTGGAAAA	TCAGTGTGCT	[419]
20_Euglena_clara	GCTGCGGTAA	TTCCAGCTCC	AAGAGCGTAC	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGCAG	CGATCGACAC	CCTGAAAA	TCAGTGTGCT	[496]
21_Discoplastis_adunca	GCTGCGGTAA	TTCCAGCTCC	GCTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TCGCTGTAC	TGCCGGGTGC	TCGCTGTAC	CCTGAAAA	TCAGTGTGCT	[492]
22_Euglena_deses	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCGAAGAC	ACTTGTAGTC	TGCCGGGTGC	TCACAGACAC	CGTGGAAAA	TCAGTGTGCT	[492]
23_Euglena_geniculata	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTATCATTG	CTGCTGAAAC	ACCTGTAGTC	TGCCGGGTGC	TCAGCGACAC	CCTGAAAA	TCAGTGTGCT	[500]
24_Euglena_granulata	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACCTGTAGTC	TGCCGGGTTG	TTATCGACAC	TGTGAAAA	CAAGCGTGCC	[491]
25_Euglena_hiemalis	GCTGCGGTAA	TTCCAGCTCC	AAGAGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCTGGGCTG	TCAGTGACAC	CCTGAAAA	TCAGTTTGCT	[500]
26_Eutreptia_pertyi	GCTGCGGTAA	TTCCAGCTCC	AAGTGCGTAT	ACTAATGTTG	CTGCTGAAAC	GCTTGTAGTC	CGACTGGTGC	CATGTGTAC	TGTGAACAA	CCAATGTGAC	[475]
27_Euglena_stellata	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTATCATTG	CTGCTGAAAC	ACCTGTAGTC	TGCCGGGTTG	TCGTTGAGAT	CCTGAAAA	TCAGTGTGCT	[479]
28_Euglena_longa	GCTGCGGTAA	TTCCAGCTCC	AAGGNCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGCTG	TCAGTGACAC	CCTGAAAA	TCAGTGTGCT	[464]
29_Euglena_viridis	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	GCTTGTAGTC	TGCCGGGTTG	TTGCTGAAAC	CCTGAAAA	TCAGTGTGCT	[497]
30_Euglena_exilis	GCTGCGGTAA	TTCCAGCTCC	AGCGGCGTAT	ACTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCGGGGCC	TCGCTGTAC	CGTGGAAAA	CAAGTGTGCC	[498]
31_Euglena_gracilis	GCTGCGGTAA	TTCCAGCTCC	AAGAGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCTGGGCTG	TCAGTGACAC	CCTGAAAA	TCAGTTTGCT	[464]
32_Lepocinclis_steinii	GCTGCGGTAA	TTCCAGCTCC	ATGTGCGTAT	ACTAACATTG	CTGTTGAAAC	ACTTGTAGTC	TGATGGGTTG	CCTCAGACAC	CGTGAAC - AC	TCAGTGTGCC	[459]
33_Lepocinclis_acus_var_major	GCTGCGGTAA	TTCCAGCTCC	AAGCGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGATGGTTGC	TCCCTGTAC	CGTGGAAAA	TCAGTGTGCT	[498]
34_Lepocinclis_tripteris	GCTGCGGTAA	TTCCAGCTCC	AAGAGCGTAT	ACTGATGTTG	CTGCTGAAAC	ACTTGTAGTC	TGATGGGT - T	TTGTTGACAC	CCTGGAGAAC	TCAGTGTGCC	[419]
35_Lepocinclis_spirogyroides	GCTGCGGTAA	TTCCAGCTCC	GAGTGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGATGGGTTG	TCTCTGACAC	CGTGGAAAA	TCAGTGTGCC	[420]
36_Lepocinclis_ovum	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	GCTTGTAGTC	TGATGGGTCG	TCACTGCCAC	CGTGGAAAA	TCAGTGTGCT	[498]
37_Monomorphina_aenigmatica	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAATATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGTGC	TTGCACTAC	CCTGAAAA	TCAGTGTGCC	[497]
38_Colacium_mucronatum	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	GCTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCAGGAGC	TTGAAGTCAC	TATGGATAAC	TCAGTGTGCT	[474]
39_Monomorphina_sp	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGTTG	TCGCACTAC	CGTGGAAAA	TCAGTGTGCT	[500]
40_Peranema_trichophorum	GCTGCGGTAA	TTCCAGCTCC	AGTAGCGTAT	ACTAATATTG	TTGCTGAAAC	GCTTCGTAGTT	CGATGAGTTG	CGAAAGTCAT	TGTGAACAAA	TCAGTGTGAT	[496]
41_Phacus_orbicularis	CCCGCGGTAA	TTCCAGCTCC	AAGAGTGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	GGACGGGTGC	TCAATGTAC	CGTGGAAAA	TCAGTGTGCT	[461]
42_Phacus_similis	GCTGCGGTAA	TTCCAGCTCC	AAGAGTGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	GGATGGGTGT	GTGCTGTAC	CCTGAAAA	TCAGTGTGCT	[462]
43_Monomorphina_pyrum	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGTTG	TTGCACTAC	CGTGGAAAA	TCAGTGTGCT	[499]
44_Strombomonas_acuminata	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCGGGTAC	TCACTGACAC	CCTGAAAA	TCAGTGTGCT	[420]
45_Strombomonas_napiformis	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCGGGCAC	TCACTGACAC	CCTGAAAA	TCAGTGTGCT	[499]
46_Strombomonas_borysteniensis	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCGGGTGC	CGCTTGACAC	CCTGAAAA	TCAGTGTGCT	[500]
47_Strombomonas_ovalis	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCGGGTGC	CGCTTGACAC	CCTGAAAA	TCAGTGTGCT	[499]
48_Trachelomonas_echinata	GCTGCGGTAA	TTCCAGCTCC	ATTGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGTGC	TCAGAGCAAC	CCTGGTAAAC	TCAGTGTGCT	[499]
49_Strombomonas_triquetra	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCGGAAAC	ACTTGTAGTC	TGCCGGGTTG	TCACTGACAC	CCTGAAAA	TCAGTGTGCT	[500]
50_Eutreptiella_pomquetensis	GCTGCGGTAA	TTCCAGCTCC	AATGGCGTAT	ACTAATATTG	CTGCTGAAAC	GCTTGTAGTC	GACCGCGTGC	CGGCTGTAC	CGTGAACAA	TCAGAGTGTG	[463]
51_Trachelomonas_lefevrei	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGTGC	TCGCTGTAC	CGTGGAAAA	TCAGTGTGCT	[499]
52_Trachelomonas_planctonica	-CTGCGGTAA	TT - CAGCTCC	AAGGGCGTAT	ACTAACATTG	-CGCTGAAAC	ACTTGTAGTC	TGCCGGGTGC	TCGCTGTAC	CCTGAAAA	TCAGTGTGCT	[494]
53_Trachelomonas_similis	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGCAG	TCAGTGACAC	CCTGAAAA	TCAGTGTGCT	[499]
54_Trachelomonas_volvocina	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAATATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGTGC	TCAGTGACAC	CCTGAAAA	TCAGTGTGCT	[463]
55_Trachelomonas_conradi	GCTGCGGTAA	TTCCAGCTCC	AAGGGCGTAT	ACTAACATTG	CTGCTGAAAC	ACTTGTAGTC	TGCCGGGTGC	TCAGTGTAC	CCTGAAAA	TCAGTGTGCC	[500]

	510	520	530	540	550	560	570	580	590	600	
01_Rhynchomonas_nasuta	CaAaGcAgTc	AtTaGaAaGc	AtGgGaTaAc	AaAgCaGcAg	CcTaTgC-Tt	-tTgAtTaCg	GgTaAgGaGg	GtCgGgGgAa	AaCgTaCtTg	GgCgTgAgAg	[595]
02_Blastocrithidia_culicis	CAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAGAGA-CAG	CCTTGGC-TT	-TTGATTATG	GGCAAGGAGG	GTGCGGGGAG	AACGTACTCG	AGCGTCAGAG	[583]
03_Herpetomonas_muscarum	TAAAGCAGTC	ATTAGCAAGC	ATGGGATAAC	AAAGGAGCAG	CCTTTAC-TT	-TTGATTATG	AGAAAGGAGG	GTGCGGGGAG	AACGTACTGG	GGCGTGAGAG	[582]
04_Leishmania_braziliensis	TAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAAGGAGCAG	CCTCTAC-TT	-TTGATTATG	GAGAAGGAGG	GTGCGGGGAG	AACGTACTGG	GGCGTCAGAG	[588]
05_Leptomonas_bifurcata	TAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAAGGAGCAG	CCTCTAC-TT	-TTGATTATG	GAGAAGGAGG	GTGCGGGGAG	AACGTACTGG	GGCGTCAGAG	[582]
06_Leishmania_infantum	TAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAAGGAGCAG	CCTCTAC-TT	-TTGATTATG	GAGAAGGAGG	GTGCGGGGAG	AACGTACTGG	GGCGTCAGAG	[577]
07_Leishmania_major	TAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAAGGAGCAG	CCTCTAC-TT	-TTGATTATG	GAGAAGGAGG	GTGCGGGGAG	AACGTACTGG	GGCGTCAGAG	[588]
08_Leishmania_tarentolae	TAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAAGGAGCAG	CCTCTAC-TT	-TTGATTATG	GAGAAGGAGG	GTGCGGGGAG	AACGTACTGG	GGCGTCAGAG	[594]
09_Trypanosoma_cruzi	CAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAAGGAGCAG	CCTCTGC-TT	-TTGATTATG	GGGAAGGAGG	GTGCGGGGAG	AACGTACTGG	TGCGTCAGAG	[587]
10_Trypanosoma_brucei	CAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	GAAGCATCAG	CCCTGGC-TT	-TTGATTATG	GGGAAGGAGG	GTGCGGGGAG	AACGTACTGG	TGCGTCAGAG	[596]
11_Bodo_saltans	CAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAAGGAGCAA	CCTACGC-TT	-TTGATTATG	GTTAAGGAGG	GTGCGGGGAA	AACGTACTGG	ATCGTGAGAG	[593]
12_Dimastigella_trypaniformis	CAAAGCAGTC	ATTAGAAAGC	ATGGGATAAC	AAAGGAGCAG	CCTATTC-TT	-TTGATTATG	GACAAGGAGG	GTGCGGGGAA	AACGTACTGG	AGGTGTAGAG	[588]
13_Rhynchobodo_sp	CAAAGCAGTC	ATTAGAAAGC	ACGGGATAAC	AAAGGAG---	CCTACGC-TT	-TTGATTATG	GATAAGGAGG	GTGCGGGGAG	AACGTACTGG	AGCGTGAGAG	[571]
14_Colacium_elongatum	CAAAGCATTG	ATGTCCATCC	ATGGAATGTC	AAAGCATCGA	CCAAGTTGTG	CTGGAATGTG	GTGGATGGGA	CTCAGAGTGA	AAGATATTGC	AACGCCAGAG	[598]
15_Ploeotia_oblonga	CAAACAGGCG	ATGATCTAGC	ATGGGAATGAC	AGGGGAGTGA	CCTCTGGCG	TTGGACCGCG	GTCGAATCGG	GACGGTGAGA	TACGTATTGC	AGCGCTAGTG	[595]
16_Cryptoglena_pigra	CAAAGCATCC	ATGTCCATCC	ATGGAACGTC	ACAGCATTGA	CCAGATC--C	-TTC-----	---GATGGAA	CTCAGGGTCA	AAGATATTGC	AGCGCCAGAG	[551]
17_Cryptoglena_skujai	CAAAGCATCC	ATGTCCATCC	ATGGAACGTC	ACAGCATTGA	CCAGATC--C	-TTC-----	---GATGGAA	CTCAGGGTCA	AAGATATTGC	AGCGCCAGAG	[551]
18_Colacium_sp	CAAAGCATCC	ATGTCCATCC	ATGGAATGTC	AAAGCATCGA	CCTTGTGTG	CTGGAATGTG	GTGGATGGGA	CTCAGAGTGA	AAGGTATTGC	AGCGCTAGAG	[598]
19_Colacium_vesiculosum	CAAAGCATCC	ATGTCCATCC	ATGGAATGTC	AAAGCATCGA	CCAAGTTGTG	CTGGAATGTG	GTGGATGGGA	TTGAGAGTGA	AAGATATTGC	AACGCTAGAG	[519]
20_Euglena_clara	CAAAGCAACC	ATGCCTGTCC	ATGGAATGTC	GCAGCATCGA	CCGAGTTGTG	CGTCCCGCAC	ATTGAGGGGT	CTCGAGGTGG	ACGATATTGT	AGCGCCAGAG	[596]
21_Discoplastis_adunca	CAAAGCATCC	ATGTCCATCC	ATGGAATGTC	AAAGCATCGA	CCTTGTGTG	CTGCAATGTG	AGGTATTGGA	CTCGGGGTGA	ACGATATTGC	AGCGCTAGAG	[592]
22_Euglena_deses	CAA--GCCTCC	ATGTCTCTCC	ATGGCATGTC	AAGGCACCGA	CCAAGTC--C	-TTGCACCGG	CGCGATGGGA	CTCAGGGTGA	AAGGTACTAC	CGCGCCAGAG	[588]
23_Euglena_geniculata	CAAAGCGTCC	ATGTCTCTCC	ATGGAACGTC	AAAGCATCGA	CCGAGCTGTG	CTGCAATGTG	GTGGATGGGA	CTCGGGGTGA	AAGGAACGTC	AGCGCCAGAG	[600]
24_Euglena_granulata	CAA--GCAGCC	ATGCTCCGTC	ATGGAATGTC	AGAGCATCGA	CCGAACC-CG	CCGGTGCGTG	GGGGAGGGGA	CTCAGGGTGA	AAGATACTGC	AACGCCAGAG	[589]
25_Euglena_hiemalis	CAAAGAAACC	ATGTCCATCC	ATGGAATGTC	AAGGCATCGA	CCGAGTC--C	-TTG-----	---GATGGGA	CTCGGGGTGA	AAGATATGGG	AGCGCCAGAG	[588]
26_Eutreptia_pertyi	CAAAGCAACT	T-GATTCTCT	ACGAGATGCG	AATGGATCGA	C-----TGCA	-TG-----A	ATGGATGGGT	ACCGAAGTGC	TGGATATTGA	CACGCCAGAG	[562]
27_Euglena_stellata	CAAGGCGTCC	ATGTCCATCC	ATGGAATGTC	CAAGCAACGA	C-----C-CT	-TG-----A	CGTGATGGGA	TTGAGGGTGA	AAGATACTGC	AACGCCAGAG	[566]
28_Euglena_longa	CAAAGCATCC	ATGTCCATCC	ATGGAACGTC	AAGGCATCGA	CCGAGTC--C	-TTG-----	---GATGGGA	CTCGGGGTGA	AAGATATGGG	AGCGCCAGAG	[552]
29_Euglena_viridis	CAAAGCATCC	ATGTCTATCT	ATGGAATTTT	AAAGCATCGA	CCGAGTTGTG	CTGCAATGTG	GTGGATGGGA	TTGAGGGTGA	AAGGTATTGC	AGCGCCAGAG	[597]
30_Euglena_exilis	CAAAGCAACC	ATGTCTCTCC	GTGGTATGTC	GTGCAACCGA	CTGGATG--C	-TGCAACACG	GGGAGGGGGA	CTCAGGGTGC	AAGGAACGTC	AACGCCAGAG	[595]
31_Euglena_gracilis	CAAAGAAACC	ATGTCCATCC	ATGGAATGTC	AAGGCATCGA	CCGAGTC--C	-TTG-----	---GATGGGA	CTCGGGGTGA	AAGATATGGG	AGCGCCAGAG	[552]
32_Lepocinclis_steinii	CAAAGCATCC	ATGTCCGTC	ATGGAATGTC	AAAGCATCGA	CTGGTTGTG	CTG-----A	ATGGATGGGA	CTCGGGGTGA	AAGGTATTGC	AGCGCCAGAG	[553]
33_Lepocinclis_acus_var_major	CAAGGCGTCC	ATGCCCCGTC	ATGGAATGTC	GCAGTATCGA	CCTGATTGTG	CTGG-----T	CCCGATGGGA	CTCAGGGTGA	AAGATATTGC	AACGCCAGAG	[593]
34_Lepocinclis_tripteris	CAAAGCATCC	ATGCACGTCC	ATGGAATGTC	ATAGCATTGA	CCAGATTGTG	CTGG-----	---GAGGGGA	CTCGGGGTGG	AAGGTATTGC	AACGCCAGAG	[510]
35_Lepocinclis_spirogyroides	CAAAGCAGCC	ATGCCCCGTC	ATGGAATGTC	AAAGCATCGA	CCACCTTGTG	CTGG-----	---GATGGGA	CTCAGGGTGG	AAGGTATTGC	AACGCCAGAG	[511]
36_Lepocinclis_ovum	CAAGGCAACC	ATGTTCGTTT	ATGGCATGGC	AAAGCATCGA	TCAGATTGTG	CTGG-----	---GAGGGGA	CTCAGGGTGA	AAGATATTGC	AACGCCAGAG	[589]
37_Monomorphina_aenigmatica	CAAAGCATCC	ATGCCTGTCC	ATGGAATGTC	AAAGCAACGA	CCGAGTTGTG	CTGGAATCTG	GTGGATGGGA	CTCAGAGTGC	AAGATATTGC	AACGCCAGAG	[597]
38_Colacium_mucronatum	CAAAGCATTC	ATGTCCATCC	ATGGAATGTC	AAAGCATCGA	CCAAGTTGTG	CTGGAATGTG	GTGGATGGGA	CTCAGAGTGA	AAGATATTGC	AACGCTAGAG	[574]
39_Monomorphina_sp	CAAAGCATCC	ATGTCCATCC	GTGGAATGAC	AAAGCACC GC	CCAGAT----	-----	---GATGGGA	CTCAAGGTGG	AAGATATTGC	AGCGCCAGAG	[583]
40_Peranema_trichophorum	CAAAGCAGCT	ATGCTTCAGC	ATGGAATGAC	AAGGGATCGA	C-----T-TC	-TGCAATGTG	GGGAATGGGA	GTGAGGTGG	TGTGTATTTC	TGCGTTAGAG	[589]
41_Phacus_orbicularis	CAAAGCATCC	ATGCCCCGTC	ATGGAATGTC	AAAGCATCGA	CCTGATTGTG	CTGGGTCCTG	GGCGATGGGA	CTCGGGGTGA	AGGATATGGC	TGCGCCAGAG	[561]
42_Phacus_similis	CAAAGCAGCC	ATGCCCCGTC	ATGGAATGTC	AAAGTATCGA	CCTGATTGTG	CTGCAATGTG	GTGGATGGGA	CTCTGGGTGA	AGGGTATTGT	AGCGCCAGAG	[562]
43_Monomorphina_pyrum	CAAAGCATCC	ATGTCCATCC	GTGGAATGAC	AAAGCACC GC	CCAGAT----	-----	---GATGGGA	CGCAAGGTGG	AAGGTATTGC	AACGCCAGAG	[582]
44_Strombomonas_acuminata	CAAAGCGTCC	ATGCCTGTCC	ATGGAATGTC	AAAGCATCGA	CCGAATGTG	CAGGAATGTG	GTGGATGGGA	CTCAGGGTGG	AAGGTATTGG	AGCGCCAGAG	[520]
45_Strombomonas_napiformis	CAAAGCGTCC	ATGCCTGTCC	ATGGAATGTC	AAAGCATCGA	CCGAATGTG	GTGGATGGGA	GTGGATGGGA	CTCAGGGTGG	AAGGTATTGG	AGCGCCAGAG	[599]
46_Strombomonas_borysteniensis	CAAAGCGTCC	ATGCTCATCC	ATGGAATGTC	AAAGCATCGA	CCGAATGTG	CAGGAATGTG	GTGGATGGGA	CTCAGGGTGG	AAGGTATTGG	AGCGCCAGAG	[600]
47_Strombomonas_ovalis	CAAAGCATCC	ATGCTCATCC	ATGGAATGTC	AAAGCATCGA	CCGAATGTG	CAGGAATGTG	GTGGATGGGA	CTCAGGGTGG	AAGGTATTGG	AGCGCCAGAG	[599]
48_Trachelomonas_lefevrei	TCAAGCAACC	ATGCC--ATCC	AGTGAATTGC	AAGGCAACGA	TGACGCTGTG	CTGCAATGTG	GGGGATGGAA	GTGAGAGTGG	AAGATACTGC	AGCGCCAGAG	[598]
49_Strombomonas_triquetra	CAAAGCATCC	ATGCCTGTCC	ATGGAATGTC	AAAGCATCGA	CCAACCTGTG	CAGGAATGTG	GTGGATGGGA	CTCAGGGTGG	AAGGTATTGG	AGCGCCAGAG	[600]
50_Eutreptiella_pomquetensis	CAAAACGACC	ATGCTTCCGC	ATGGGATGAC	AAAGGATCGA	CCAGATTTTG	CTGCAATGCG	CGCGACGGGA	GTGCTGGTGA	TGGGTATTGA	CGCGTAGAG	[563]
51_Trachelomonas_lefevrei	CAAAGCATCC	ATGTCCGTC	ATGGAATGTC	AAAGCATCGA	CGGGTC--C	-TTG-----	---GATGGGA	CTCAGGGTGA	AAGGTATTGG	AGCGCCAGAG	[587]
52_Trachelomonas_planctonica	CAAAGCATCC	ATGTCCGTC	ATGGAATGTC	AAAGCATCGA	CCGAGTTGTG	CTGGAATCTG	GTGGATGGGA	CTCAGGGTGA	AAGGTATTGG	AACGCCAGAG	[594]
53_Trachelomonas_similis	CAAAGCGACC	ATGCCTATCC	ATGGAATGTC	AAAGCACC GA	CCTGATG--C	-TTC-----	---GATGGGA	TTGAGGGCGG	AAGGTACTGC	AACGCCAGAG	[587]
54_Trachelomonas_volvocina	CAAAGCGTCC	ATGCCTGTCC	ATGGAATGTC	AAAGCACC GA	CCGAGTTGTG	CTGCAATGTG	GGGGATGGGA	GTGAGGGCGA	AAGGTATTGG	AACGCCAGAG	[563]
55_Trachelomonas_conradi	CAACGCGTCC	ATGTCCGTC	ATGGAATGTC	AAAGCATCGA	CCGAGTTGTG	CTGC-----	---GATGGGA	CTCAGGGTGA	AAGGTATTGG	AACGCCAGAG	[591]

	610	620	630	640	650	660	670	680	690	700	
01_Rhynchomonas_nasuta	GtGaAaTtCt	TaGaCcGcCc	AaAgAcGaAc	TaCaGcGaAg	GcAtTtCcCa	GgAtAcCtTc	CtCaAtCaAg	AaCcAaAgTg	TgGgGaTcGa	AgAtGaTtAg	[695]
02_Blastocrithidia_culicis	GTGAAATTCT	TAGACCGATC	GAAGACGAAC	TACAGCGGAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[683]
03_Herpetomonas_muscarum	GTGAAATTCT	TAGACCGCAC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	CGGGGATCGA	AGATGATTAG	[682]
04_Leishmania_braziliensis	GTGAAATTCT	TAGACCGCAC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[688]
05_Leptomonas_bifurcata	GTGAAATTCT	TAGACCGCAC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[682]
06_Leishmania_infantum	GTGAAATTCT	TAGACCGCAC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[677]
07_Leishmania_major	GTGAAATTCT	TAGACCGCAC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[688]
08_Leishmania_tarentolae	GTGAAATTCT	TAGACCGCAC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[694]
09_Trypanosoma_cruzi	GTGAAATTCT	TAGACCGCAC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGGATCGA	AGATGATTAG	[687]
10_Trypanosoma_brucei	GTGAAATTCT	TAGACCGCAC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGGATCGA	AGATGATTAG	[696]
11_Bodo_saltans	GTGAAATTCT	TAGACCGTTC	CAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[693]
12_Dimastigella_trypaniformis	GTGAAATCCT	TAGACCACTC	AAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[688]
13_Rhynchobodo_sp	GTGAAATTCT	TAGACCGCTC	AAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	TGGAGATCGA	AGATGATTAG	[671]
14_Colacium_elongatum	GTGAAATTCT	TAGACCGTTG	CAAGATCAAC	GGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGCG	TGGGGAGCGA	AGATGATCAG	[698]
15_Ploeotia_oblonga	GTGAAATACT	TAGACCGTTC	AAAGACGAAC	TACAGCGAAG	GCATTTTCAA	GGATACCTTC	CTCAATCAAG	AACCAAAGTG	AGCGATAGCA	AGATGATCAG	[695]
16_Cryptoglana_pigra	GTGAAATTCT	TAGACCGCTG	CAGGATCAAC	AGCAGCGAAG	CGCTTGGCAA	GTGCGATGTT	ATCGATCAAG	AATGAAAGCG	TGGGGAGCAA	AGATGATCAG	[651]
17_Cryptoglana_skujai	GTGAAATTCT	TAGACCGCTG	CAGGATCAAC	AGCAGCGAAG	CGCTTGGCAA	GTGCGATGTT	ATCGATCAAG	AATGAAAGCG	TGGGGAGCAA	AGATGATCAG	[651]
18_Colacium_sp	GTGAAATTCT	TAGACCGTTC	CAAGATCAAC	GGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGCG	TGGGGAGCAA	AGATGATCAG	[698]
19_Colacium_vesiculosum	GTGAAATTCT	TAGACCGCTG	CAAGATCAAC	GGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGCG	TGGGGAGCAA	AGATGATCAG	[619]
20_Euglena_clara	GTGAAATTTT	CAGATCGCTG	CAAGATCCAC	GGCAGCGAAG	CGCTTGTCTA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGTA	TGGGGAGCAA	AGATGATCAG	[696]
21_Discoplastis_adunca	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGTG	TGGGGAGCAA	AGATGATCAG	[692]
22_Euglena_deses	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	GGCAGCGAAG	CGCTTGTCTA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGTC	CGGGGAGCAA	AGATGATCAG	[688]
23_Euglena_geniculata	GTGAAATTCT	TAGATCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	GATGAGAGTC	TGGGGAGCGA	AGATGATCAG	[700]
24_Euglena_granulata	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	GATGAGAGTC	CGGGGAGCAA	AGATGATCAG	[689]
25_Euglena_hiemalis	GTGAAATTCT	TAGATCGCTG	CCAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGTT	CGGGGAGCAA	AGATGATCAG	[688]
26_Eutreptia_pertyi	GTGAAATTCT	TAGACCGCTG	TGGGATCGAG	TGCGGCAAG	CGCTTTGCAA	GATTCGTGTC	GTTGATCAAG	AACGAGAGTG	TGGGGAGCAA	AGATGATCAG	[662]
27_Euglena_stellata	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	GATGAGAGTC	TGGGGAGCAA	AGAGGATCAG	[666]
28_Euglena_longa	GTGAAATTCT	TAGATCGCTG	CCAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGTT	CGGGGAGCAA	AGATGATCAG	[652]
29_Euglena_viridis	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	GATGAGAGTC	CGGGGAGCAA	AGATGATCAG	[697]
30_Euglena_exilis	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AGCGAGAGCC	TGGGGATCGA	AGAGGATCAG	[695]
31_Euglena_gracilis	GTGAAATTCT	TAGATCGCTG	CCAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGTT	CGGGGAGCAA	AGATGATCAG	[652]
32_Lepocinclis_steinii	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGTGAGTG	TGGGGAGCGA	AGATGATCAG	[653]
33_Lepocinclis_acus_var_major	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGTGAGTG	TGGGGAGCAA	AGATGATCAG	[693]
34_Lepocinclis_tripteris	GTGAAATTCT	CAGATCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGTGAGTG	TGGGGAGCGA	AGATGATCAG	[610]
35_Lepocinclis_spirogyroides	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	GGCAGCGGAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGTGAGTG	TGGGGAGCAA	AGATGATCAG	[611]
36_Lepocinclis_ovum	GTGAAATTCT	TAGATCGCTG	CAAGATGACC	CGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AACGTGAGTG	TGGGGAGCGA	AGATGATCAG	[689]
37_Monomorphina_aenigmatica	GTGAAATTCT	TAGACCGCTG	CACGATCCAC	AGCAGCGAAG	CGCTTTGCAA	GGATGAGTCC	GTCGATCAAG	AATGAGAGCC	TGGGGATCCA	AGATGATTAG	[697]
38_Colacium_mucronatum	GTGAAATTCT	TAGACCGTTC	CAAGATCAAC	GGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGCG	TGGGGAGCGA	AGATGATCAG	[674]
39_Monomorphina_sp	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGCC	TGGGGAGCAA	AGGTGATTAG	[683]
40_Peranema_trichophorum	GTGAAATTCT	TAGACCGTAG	AAGGACATAC	TGCGGCGAAG	CGCTTTGCAA	GTCATGCTCT	GTCATCAAG	AACGAAAGTG	TGGGGAGCAA	AGATGATCAG	[689]
41_Phacus_orbicularis	GTGAAATTCT	TAGACCGCTG	CCAGATCAAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGTG	TGGGGAGCGA	AGATGATCAG	[661]
42_Phacus_similis	GTGAAATTCT	TAGACCGCTG	CAAGATCAAC	TGCAGCGAAG	CGCTTTGCAA	GTGCGTGTTC	GTCGATCAAG	AATGAGAGTG	TGGGGAGCGA	AGATGATCAG	[662]
43_Monomorphina_pyrum	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCGGCGAAG	CGCTTTGCAA	GTTTGAAGTCC	GTCGATCAAG	AATGAGAGCC	TGGGGAGCAA	AGGTGATTAG	[681]
44_Strombomonas_acuminata	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGAGAGCT	CGGGGAGCGA	AGATGATCAG	[620]
45_Strombomonas_napiformis	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGAGAGCT	CGGGGAGCGA	AGATGATCAG	[699]
46_Strombomonas_borysteniensis	GTGAAATTCT	TAGACCGCTG	ACAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGAGAGCG	TGGGGAGCGA	AGATGATCAG	[700]
47_Strombomonas_ovalis	GTGAAATTCT	TAGACCGCTG	ACAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGAGAGCG	TGGGGAGCGA	AGATGATCAG	[699]
48_Trachelomonas_lefevrei	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGAGAGTG	TGGGGAGCGA	AGATGATCAG	[698]
49_Strombomonas_triquetra	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGAGAGCG	TGGGGAGCGA	AGATGATTAG	[700]
50_Eutreptiella_pomquetensis	GTGAAATTCT	TAGATCGCTT	TGGGACCTTC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AACGAGAGTG	TGGGGAGCGA	AGATGATCAG	[663]
51_Trachelomonas_achineta	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGAGAGCG	TGGGGAGCGA	AGATGATCAG	[687]
52_Trachelomonas_planctonica	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGARAGTG	TGGGGAGCAR	AAATGATCAG	[694]
53_Trachelomonas_similis	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGARAGTG	TGGGGAGCGA	AGATGATCAG	[687]
54_Trachelomonas_volvocina	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGARAGTG	TGGGGAGCGA	AGATGATCAG	[663]
55_Trachelomonas_conradi	GTGAAATTCT	TAGACCGCTG	CAAGATCCAC	TGCAGCGAAG	CGCTTTGCAA	GAGCATGTTT	GTCGATCAAG	AATGARAGTG	TGGGGAGCGA	AGATGATCAG	[691]

	710	720	730	740	750	760	770	780	790	800	
01_Rhynchomonas_nasuta	AgACCaTtGt	AgTcCaC-AC	CaCaAaCgGt	GaCcCcAtTc	AaC-CgCgAa	AgCgGtTaCa	GtCtCaGgGg	-gAgTaCgTt	CgCaAgAgTg	AaAcTtAaAg	[792]
02_Blastocrithidia_culicis	AGACCATTGT	AGTCCAC-AC	TGCAAAACGAT	GACACCAC-C	AAACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[781]
03_Herpetomonas_muscarum	AGACCATTGT	AGTCCGC-AC	TGCAAAACGAT	GACACC--C	TAATCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[778]
04_Leishmania_braziliensis	AGACCATTGT	AGTCCAC-AC	TGCAAAACGAT	GACACCAC--	CAACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[785]
05_Leptomonas_bifurcata	AGACCATTGT	AGTCCAT-AC	TGCAAAACGAT	GACACCAC--	CAACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[779]
06_Leishmania_infantum	AGACCATTGT	AGTCCAC-AC	TGCAAAACGAT	GACACCAC--	CAACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[774]
07_Leishmania_major	AGACCATTGT	AGTCCAC-AC	TGCAAAACGAT	GACACCAC--	CAACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[785]
08_Leishmania_tarentolae	AGACCATTGT	AGTCCAC-AC	TGCAAAACGAT	GACACCAC--	CAACCCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[791]
09_Trypanosoma_brucei	AGACCATTGT	AGTCCAC-AC	TGCAAAACGAT	GACACCT-TA	A-ATCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[784]
10_Trypanosoma_brucei	AGACCATTGT	AGTCCAC-AC	TGCAAAACGAT	GACACCTATC	AA-TCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[794]
11_Bodo_saltans	AGACCATTGT	AGTCCAC-AC	CACAAAACGGT	GACACCA--T	AA-TCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[789]
12_Dimastigella_trypaniformis	AGACCATTGT	AGTCCAC-AC	CACAAAACGGT	GACCCCATCC	AAATCGCGAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[787]
13_Rhynchobodo_sp	AGACCATTGT	AGTCCAC-AC	CACAAAACGGG	GACCC--T	TTTTCGCAAA	AGCGGTTACA	GTCTCAGGGG	GGAGTACGTT	CGCAAGAGTG	AAACTTAAAG	[767]
14_Colacium_elongatum	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCACTC	CATCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[798]
15_Ploeotia_oblonga	ATACCGTTGT	AGTCCA-TGC	TGTAACCTAT	GCCAGCTCCC	---CGAGAA	ATCGCC--TG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[788]
16_Cryptoglana_pigra	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCGCAC	CCACCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GAAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[751]
17_Cryptoglana_skujai	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCGCAC	CCACCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GAAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[751]
18_Colacium_sp	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCACTC	A---CGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[795]
19_Colacium_vesiculosum	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCACTC	TATTCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[719]
20_Euglena_clara	ACACCGTCGT	AGTCCATCAC	CGTAAACGAT	GCCGGC-CAA	GA--CGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[793]
21_Discoplastis_adunca	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCACTC	GCCGGCAGAG	ATCGCCTTTG	GGTTCCGGGG	GAAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[792]
22_Euglena_deses	ACACCGTCGT	AGTCCGC-AC	TGTAACCTAT	GCCAGCACTC	A---CGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[784]
23_Euglena_geniculata	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCGACC	CAGCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[800]
24_Euglena_granulata	ACACCGTCGT	AGTCCGC-AC	CGTAAACCTAT	GCCAGCACTC	GCCAGCCAAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[788]
25_Euglena_hiemalis	ACACCGTCGT	AGTCCGCCAC	TGTAACCTAT	GCCGGCTCAG	A---CGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[784]
26_Eutreptia_pertyi	ATACCGTTGT	AGTCCATCTC	CGTAAACGAT	GTGAGCTCCA	CCATCGAGAA	ATCGCCTTTG	GATTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[762]
27_Euglena_stellata	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCGCAC	TAG-CGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[765]
28_Euglena_longa	ACACCGTCGT	AGTCCGCCAC	TGTAACCTAT	GCCGGCTCAG	A---CGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[749]
29_Euglena_viridis	ACACCGTCGT	AGTCCGAT-AC	TGTAACCTAT	GCCAGCACCC	AGCCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[796]
30_Euglena_exilis	ACACCGTCGT	AGTCCAGCGC	CGTAAACCTAT	GCCGGCCCAA	CCCGCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[795]
31_Euglena_gracilis	ACACCGTCGT	AGTCCGCCAC	TGTAACCTAT	GCCGGCTCAG	A---CGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[749]
32_Lepocinclis_steinii	ACACCGTCGT	AGTCCACAAC	TGTAACCTAT	GCCAGCACCA	CATCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	TGCAAGGCTG	AAACTTAAAG	[753]
33_Lepocinclis_acus_var_major	ACACCGTTGT	AGTCCACAAC	TGTAACCTAT	GCCAGCACCA	A---CGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[793]
34_Lepocinclis_tripteris	ACACCGTCGT	AGTCCACAAC	TGTAACCTAT	GCCAGCCACA	CCCCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[710]
35_Lepocinclis_spirogyroides	ACACCGTCGT	AGTCCACGAC	TGTAACCTAT	GCCGGCACGC	CATCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[711]
36_Lepocinclis_ovum	ACACCGTCGT	AGTCCACAAC	TGTAACCTAT	GCCAGCACAC	CATCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[789]
37_Monomorphina_aenigmatica	ACACCGTCGT	AGTCCAGAGC	TGTAACCTAT	GCCAGCCCA	TTGGCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[797]
38_Colacium_mucronatum	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCACTC	CATTTCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[774]
39_Monomorphina_sp	ACACCTCCGT	AGTCCAGAGC	CGTAAACCTAT	GCCGGCACCA	CCACCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	TGCAAGGCTG	AAACTTAAAG	[783]
40_Peranema_trichophorum	ATACCGTTGT	AGTCCATCTC	TGTAACCTAT	GTCA-GTTAC	CGGACGAGAA	ATCACCTATG	GATTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[788]
41_Phacus_orbicularis	ACACCGTCGT	AGTCCACAAC	TGTAACCTAT	GCCAGCCCTT	CGAGCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[761]
42_Phacus_similis	ACACCGTCGT	AGTCCACAAC	TGTAACCTAT	GCCAGCCCA-	G---CGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[758]
43_Monomorphina_pyrum	ACACCTCCGT	AGTCCAGAGC	TGTAACCTAT	GCCGGCACCA	CCACCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	TGCAAGGCTG	AAACTTAAAG	[781]
44_Strombomonas_acuminata	ACACCATCGT	AGTCCGAAGC	TGTAACCTAT	GCCAGCACAA	CTCTCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[720]
45_Strombomonas_napiformis	ACACCATCGT	AGTCCGAAGC	TGTAACCTAT	GCCAGCACAA	CCTTCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[799]
46_Strombomonas_borysteniensis	ACACCATCGT	AGTCCATGTC	TGTAACCTAT	GCCAGC--T	CATCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[797]
47_Strombomonas_ovalis	ACACCATCGT	AGTCCATGTC	TGTAACCTAT	GCCAGCCATC	AG--CGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[797]
48_Trachelomonas_efeveirei	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCACTT	GCCAGCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[798]
49_Strombomonas_triquetra	ACACCATCGT	AGTCCATGTC	TGTAACCTAT	GCCAGCACCT	CATCCGAGAA	ATCGCCTTTG	GGTTCCGGGG	GGAGTACGTT	CGCAAGGCTG	AAACTTAAAG	[800]
50_Eutreptiella_pomquetensis	ATACCGTCGT	AGTCCACTAC	CGTAAACCTAT	GTCGGCCCAA	GGACCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[763]
51_Trachelomonas_lefevrei	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCTGAA	CTCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[787]
52_Trachelomonas_planctonica	ACACCGTCGT	AGTCCACAGC	TGTAACCTAT	GCCAGCACCT	CAGCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[794]
53_Trachelomonas_similis	ACACCGTCGT	AGTCCACAAC	TGTAACCTAT	GCCAGCGAAT	CATCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[787]
54_Trachelomonas_volvocina	ACACCGTCGT	AGTCCACAAC	TGTAACCTAT	GCCAGCACCC	TATCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	CGCAAGGCTG	AAACTTAAAG	[763]
55_Trachelomonas_conradi	ACACCGTCGT	AGTCCRCCTG	TGTAACCTAT	GCCAGCACCC	CACCCGAGAA	ATCGCCTTTG	GGTTCCAGGGG	GGAGTACTGT	AGCAAGGCTG	AAACTTAAAG	[791]

	810	820	830	840	850	860	870	880	890	900	
01_Rhynchomonas_nasuta	AaAtTgAcGg	AaTgGcAcCa	CaAgAtGtGa	AgCgTgCgGt	TtAaTtTgAc	TcAaCaCgGg	GaAcTtTaCc	AgAtCcGgAa	AgGgTaAgGa	TtGaCaGaTt	[892]
02_Blastocrithidia_culicis	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGGGGAGGA	TTGACAGATT	[881]
03_Herpetomonas_muscarum	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGATGAGGA	TTGACAGATT	[878]
04_Leishmania_braziliensis	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGATGAGGA	TTGACAGATT	[885]
05_Leptomonas_bifurcata	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGATGAGGA	TTGACAGATT	[879]
06_Leishmania_infantum	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGATGAGGA	TTGACAGATT	[874]
07_Leishmania_major	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGATGAGGA	TTGACAGATT	[885]
08_Leishmania_tarentolae	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGATGAGGA	TTGACAGATT	[891]
09_Trypanosoma_cruzi	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGGTGAGGA	TTGACAGATT	[884]
10_Trypanosoma_brucei	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGGTGAGGA	TTGACAGATT	[894]
11_Bodo_saltans	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGGTGAGGA	TTGACAGATT	[889]
12_Dimastigella_trypaniformis	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGGTGAGGA	TTGACAGATT	[887]
13_Rhynchobodo_sp	AAATTGACGG	AATGGCACCA	CAAGACGTGG	AGCGTGCGGT	TTAATTTGAC	TCAACACGGG	GAACTTTACC	AGATCCGGAC	AGGGTGAGGA	TTGACAGATT	[867]
14_Colacium_elongatum	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[898]
15_Ploeotia_oblonga	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[888]
16_Cryptoglena_pigra	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[851]
17_Cryptoglena_skujai	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[851]
18_Colacium_sp	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[895]
19_Colacium_vesiculosum	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[819]
20_Euglena_clara	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[893]
21_Discoplastis_adunca	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[892]
22_Euglena_deses	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ACCGCCGGGA	TGGACAGATT	[884]
23_Euglena_geniculata	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ACAGCCGGGA	TGGACAGATT	[900]
24_Euglena_granulata	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ACAGCCGGGA	TGGACAGATT	[888]
25_Euglena_hiemalis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	CAACTGGGA	TTGACAGATT	[884]
26_Eutreptia_pertyi	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTGGCTGGGA	TTGACAGATT	[862]
27_Euglena_stellata	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ACAGCTAGGA	TGGACAGATT	[865]
28_Euglena_longa	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	CAACTGGGA	TTGACAGATT	[849]
29_Euglena_viridis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	CAAGCTGGGA	TGGACAGATT	[896]
30_Euglena_exilis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[895]
31_Euglena_gracilis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	CAACTGGGA	TTGACAGATT	[849]
32_Lepocinclis_steinii	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[853]
33_Lepocinclis_acus_var_major	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[893]
34_Lepocinclis_triapteris	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[810]
35_Lepocinclis_spirogyroides	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[811]
36_Lepocinclis_ovum	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[889]
37_Monomorphina_aenigmatica	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[897]
38_Colacium_mucronatum	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[874]
39_Monomorphina_sp	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ATGGCTGGGA	TGGACAGATT	[883]
40_Peranema_trichophorum	GAATTGACGG	AGTGACACCA	CAAGGTGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAACCTTACC	AGGTGAGGAC	GCAGTTAGGA	TTGACAGATT	[888]
41_Phacus_orbicularis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GATGCTGGGA	TGGACAGATT	[861]
42_Phacus_similis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ACTGCCGGGA	TGGACAGATT	[858]
43_Monomorphina_pyrum	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAAGGTTACC	AGGTGAGGAC	GTGGCTGGGA	TGGACAGATT	[881]
44_Strombomonas_acuminata	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ATTGCCAGGA	TGGACAGATT	[820]
45_Strombomonas_napiformis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ATTGCCAGGA	TGGACAGATT	[899]
46_Strombomonas_borysteniensis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ATTGCCAGGA	TGGACAGATT	[897]
47_Strombomonas_ovalis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ATTGCCAGGA	TGGACAGATT	[897]
48_Trachelomonas_lefevrei	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ATTGCCAGGA	TGGACAGATT	[898]
49_Strombomonas_triquetra	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	ATTGCCAGGA	TGGACAGATT	[900]
50_Eutreptiella_pomquetensis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTAGCTGGGA	TTGACAGATT	[863]
51_Trachelomonas_jeffreyi	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	STTGCTGGGA	TGGACAGATT	[887]
52_Trachelomonas_planctonica	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[894]
53_Trachelomonas_similis	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	CTTGCCAGGA	TGGACAGATT	[887]
54_Trachelomonas_volvocina	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	CTTGCCGGGA	TGGACAGATT	[863]
55_Trachelomonas_conradi	GAATTGACGG	AATGGCACCA	CAAGCGGTGG	AGTATGCGGC	TTAATTTGAC	TCAACACGGG	GAATGTTACC	AGGTGAGGAC	GTTGCTGGGA	TGGACAGATT	[891]

	910	920	930	940	950	960	970	980	990	1000	
01_Rhynchomonas_nasuta	GaTtGtTcTt	TcTcGaTtCt	CtCaAtGgTg	GtGcAtGgCc	GcTtTtGgTc	GgTgGaGtGa	TtTgTtTgGt	TgAtTcCgTc	AaCgGaCgAg	AtCcAaGcTg	[992]
02_Blastocrithidia_culicis	GAGTGTTCCT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[981]
03_Herpetomonas_muscarum	GAGTGTTCCT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[978]
04_Leishmania_braziliensis	GAGTGTTCCT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[985]
05_Leptomonas_bifurcata	GAGTGTTCCT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[979]
06_Leishmania_infantum	GAGTGTTCCT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[974]
07_Leishmania_major	GAGTGTTCCT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[985]
08_Leishmania_tarentolae	GAGTGTTCCT	TCTCGATTCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[991]
09_Trypanosoma_brucei	GAGTGTTCCT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[984]
10_Bodo_saltans	GAGTGTTCCT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[994]
11_Dimastigella_trypaniformis	GATTGTTCCT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[989]
12_Rhynchobodo_sp	GAGTGTTCCT	TCTCGATCCC	CTGAATGGTG	GTGCATGGCC	GCTTTTGGTC	GGTGGAGTGA	TTTGTTTGGT	TGATTCCGTC	AACGGACGAG	ATCCAAGCTG	[987]
13_Colacium_elongatum	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[998]
14_Ploetia_oblonga	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[951]
15_Cryptoglena_pigra	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[951]
16_Cryptoglena_skujai	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[951]
17_Colacium_sp	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[995]
18_Colacium_vesiculosum	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[919]
19_Euglena_clara	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[993]
20_Discoplastis_adunca	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[992]
21_Euglena_deses	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[984]
22_Euglena_geniculata	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[1000]
23_Euglena_granulata	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[988]
24_Euglena_hiemalis	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[984]
25_Eutrepia_pertyi	GAAAGCCCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[962]
26_Euglena_stellata	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[965]
27_Euglena_longa	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[949]
28_Euglena_viridis	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[996]
29_Euglena_exilis	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[995]
30_Euglena_gracilis	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[949]
31_Lepocinlis_steinii	GAAAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[953]
32_Lepocinlis_acus_var_major	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[993]
33_Lepocinlis_triapteris	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[910]
34_Lepocinlis_spirogyroides	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[911]
35_Lepocinlis_ovum	GAAAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[989]
36_Monomorphina_aenigmatica	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[997]
37_Colacium_mucronatum	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[974]
38_Monomorphina_sp	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[983]
39_Peranema_trichophorum	GTGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[988]
40_Phacus_orbicularis	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[961]
41_Phacus_similis	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[958]
42_Monomorphina_pyrum	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ACATCTGCCT	[981]
43_Strombomonas_acuminata	GCTAGCTCTT	GCTTGATCCG	ATGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[920]
44_Strombomonas_napiformis	GCTAGCTCTT	GCTTGATCCG	ATGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[999]
45_Strombomonas_borysteniensis	GCTAGCTCTT	GCTTGATCCG	ATGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[997]
46_Strombomonas_ovalis	GCTAGCTCTT	GCTTGATCCG	ATGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[997]
47_Trachelomonas_echinata	GCTAGCTCTT	GCTTGATCCG	ATGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[997]
48_Trachelomonas_triquetra	GCTAGCTCTT	GCTTGATCCG	ATGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[998]
49_Eutrepia_pomquetensis	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[1000]
50_Trachelomonas_lefevrei	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[963]
51_Trachelomonas_planctonica	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[987]
52_Trachelomonas_similis	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[994]
53_Trachelomonas_volvocina	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[987]
54_Trachelomonas_conradi	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[963]
55_Trachelomonas_conradi	GAGAGCTCTT	TCTTGATCCT	TCGGATGGTG	GTGCATGGCC	GCTCCTGATT	GGTGGAGTGA	TTTGTCTGGT	TGATTCCGAT	AACGAGTGAG	ATCTCTGCCT	[991]

	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	
01_Rhynchomonas nasuta	CcCaGtGcCt	GaAtCgTcCt	TcTcTgCaGg	AtCcTgGtGg	TgAaAtTtTg	GgCaCaAgCa	GgTcTgTgAt	GcTcCtCaAt	GcTcTgG-Cg	AcAcGcGcAc	[1091]
02_Blastocrithidia culicis	CCCAGTCCCC	GGATTAGGCT	TCTCTGCGGG	ATCCCTGCGG	TGAGATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1081]
03_Herpetomonas muscarum	CCCAGTTTCA	GAATTGTCT	TCTCTGCGGG	ATCCTTGTGG	TGAGATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1078]
04_Leishmania braziliensis	CCCAGTTTCA	GAATTGTCT	TCTCTGCGGG	ATCCTTGTGG	TGAAATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1085]
05_Leptomonas bifurcata	CCCAGTTTCA	GAATTGTCT	TCTCTGCGGG	ATCCTTGTGG	TGAAATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1079]
06_Leishmania infantum	CCCAGTTTCA	GAATTGTCT	TCTCTGCGGG	ATCCTTGTGG	TGAAATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1074]
07_Leishmania major	CCCAGTTTCA	GAATTGTCT	TCTCTGCGGG	ATCCTTGTGG	TGAAATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1085]
08_Leishmania tarentolae	CCCAGTTTCA	GAATTGTCT	TCTCTGCGGG	ATCCTTGTGG	TGAAATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1091]
09_Trypanosoma cruzi	CCCAGTTTCA	GAATTGCCT	TCTCTGCGGG	ATCCTTGTGG	TGAGATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCCAAT	GTTCTGGGCG	ACACGCGCAC	[1084]
10_Trypanosoma brucei	CCCAGTCCGG	GGATTGTCT	TCTCTGCGGG	ATCCTTGTGG	TGAGATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1094]
11_Bodo saltans	CCCAGTACCA	GGATTGTCT	TCTCTGCGGG	ATCCTTGTGG	TGAGATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1089]
12_Dimastigella trypaniformis	CCCAGTGACT	AAATCGTTCT	TCTCTGCGGG	ATCCTTGTGG	TGAGATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1086]
13_Rhynchobodo sp	CCCAGTACCT	GAAGTTTCT	TCTCTGCGGG	ATCCTTGTGG	TGAGATTTTG	GGCAACAGCA	GGTCTGTGAT	GCTCCTCAAT	GTTCTGGGCG	ACACGCGCAC	[1067]
14_Colacium elongatum	CCCAATTCCA	TCTCGCAGCT	TCTCTGAGGT	GCGTGTCCAC	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1098]
15_Ploeotia oblonga	CTCACTTCTG	CTTCCGAAG	CCTTAGAGGT	ACCGCGGTCC	GGAAGATTTG	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1087]
16_Cryptoglena pigra	CCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGTGTCCAC	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1051]
17_Cryptoglena skujai	CCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGTGTCCAC	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1051]
18_Colacium sp	CCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGTGTCCAC	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1095]
19_Colacium vesiculosum	CCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGTGTCCAC	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1019]
20_Euglena clara	CCCAGTCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1093]
21_Discoplastia adunca	CCCAGTACTG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1092]
22_Euglena deses	CCCAGTCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1084]
23_Euglena geniculata	CCCAATCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1100]
24_Euglena granulata	CCCAGTCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1088]
25_Euglena hiemalis	CCCAGTCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1084]
26_Eutreptia pertyi	CGA-ATGCCG	GCTCGCAGCT	TCTCGAAGGT	GCGGTGCCG	GTAAGCCAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCTAGAT	GTCTGGGCG	GCACGCGCAC	[1061]
27_Euglena stellata	CCCAGTCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1065]
28_Euglena longa	CCCAGTCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1049]
29_Euglena viridis	CCCAATCTCG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1096]
30_Euglena exilis	CCCAGTCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1095]
31_Euglena gracilis	CCCAGTCCGG	GCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1049]
32_Lepocinlis steinii	CCCAATCCAG	CCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1053]
33_Lepocinlis acus var major	CCCAGTCCGG	C-TGCGAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1092]
34_Lepocinlis tripteris	CCCAATCCAG	CCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1010]
35_Lepocinlis spirogyroides	CCCAATCCAG	CCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1011]
36_Lepocinlis ovum	CCCAGTCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1089]
37_Monomorphina aenigmatica	CCCAATCC--	-----GCT	TCTCTGAGGT	GCGAGCATTT	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1088]
38_Colacium mucronatum	CCCAATTTCA	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1074]
39_Monomorphina sp	CCCAGTCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1083]
40_Peranema trichophorum	TCCAATCTTC	ATCTCGTCT	TCTCGAAGGG	ACGTGT-GAC	TGAGGTTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCTAAAT	GTCTGGGCG	GCACGCGCAC	[1087]
41_Phacus orbicularis	CCCAATCCAG	CCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1061]
42_Phacus similis	CCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1058]
43_Monomorphina pyrum	CCCAGTCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1080]
44_Strombomonas acuminata	TCCACTCCCTG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCTAGAT	GTTCTGGGCG	GCACGCGCAC	[1020]
45_Strombomonas napiformis	TCCACTCCCTG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCTAGAT	GTTCTGGGCG	GCACGCGCAC	[1099]
46_Strombomonas borysteniensis	TCCACTCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCTAGAT	GTCTGGGCG	GCACGCGCAC	[1097]
47_Strombomonas ovalis	TCCACTCCCTG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCTAGAT	GTCTGGGCG	GCACGCGCAC	[1097]
48_Trachelomonas lefevrei	TCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAT	[1098]
49_Strombomonas triquetra	TCCACTCCCTG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCTAGAT	GTCTGGGCG	GCACGCGCAC	[1100]
50_Eutreptiella pomquetensis	CCCAATCCCG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGTA	GGTCTGTGAT	GCTCCTAGAT	GTCTGGGCG	GCACGCGCAC	[1063]
51_Trachelomonas echinata	CCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1087]
52_Trachelomonas planctonica	CCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1094]
53_Trachelomonas similis	GCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1087]
54_Trachelomonas volvocina	GCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1063]
55_Trachelomonas conradi	CCCAATCCAG	TCTCGCAGCT	TCTCTGAGGT	GCGGTGCCG	GCATGCTAGA	GCCAACAGCA	GGTCTGTGAT	GCTCCAGAT	GTCTGGGCG	GCACGCGCAC	[1091]

	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	
01_Rhynchomonas_nasuta	TaCaAtGtCa	GtGaGaAcAa	GaAgG-GcGg	GgAcAaTtGg	GaCcGaGtAt	TgCaAtTaTt	GgTcAaCgAg	GaAtAtCtCg	TaGgCgTaGc	TcAtCaAaCt	[1190]
02_Blastocrithidia_culicis	TACAATGTCA	GTGAGAACAA	G-AGG-GCGG	GGACACTTGG	GACCGAGGAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1179]
03_Herpetomonas_muscarum	TACAATGTCA	GTGAGAACAA	G-AGA-GTGG	GGAGTTTTGG	GACCGAGGAT	TGCAATTATT	GGTCAACTAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1176]
04_Leishmania_braziliensis	TACAATGTCA	GTGAGAACAA	G-AGA-GTGG	GGACACTTGG	GACCGAGGAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1183]
05_Leptomonas_bifurcata	TACAATGTCA	GTGAGAACAA	G-AGA-GTGG	GGACACTTGG	GACCGAGGAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1177]
06_Leishmania_infantum	TACAATGTCA	GTGAGAACAA	G-AGA-GTGG	GGACACTTGG	GACCGAGGAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1172]
07_Leishmania_major	TACAATGTCA	GTGAGAACAA	G-AGA-GTGG	GGACACTTGG	GACCGAGGAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1183]
08_Leishmania_tarentolae	TACAATGTCA	GTGAGAACAA	G-AGA-GTGG	GGACACTTGG	GACCGAGGAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1189]
09_Trypanosoma_cruzi	TACAATGTCA	GTGAGAACAA	G-AGA-GTGG	GAACACTTGG	AACCGAGTAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1182]
10_Trypanosoma_brucei	TACAATGTCA	GTGAGAACAA	GAAGA-GCGG	GGACACTTGG	GACCGAGTAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1193]
11_Bodo_saltans	TACAATGTCA	GTGAGAACAA	G-AGA-GGGG	GGACAATTGG	GACCGAGTAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1187]
12_Dimastigella_trypaniformis	TACAATGTCA	GTGAGAACAA	GAAGG-GCGG	GGACAATTGG	GACCGAGGAT	TGCAATTATT	GGTCAACGAG	GAATGTCTCG	TAGGCGCAGC	TCATCAAACCT	[1185]
13_Rhynchobodo_sp	TACAATGTCA	GCGAGAACAA	GAAGA-GGGG	GGACAATTGG	GACCGAGGAT	TGCAATTATT	GGTCAACGAG	GAATGCCTCG	TAGGCGCAGC	TCATCAAACCT	[1166]
14_Colacium_elongatum	TACATTGACA	CAGGGAAGGA	GTGTGAATGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCAAG	TCATCAGCTT	[1198]
15_Ploeotia_oblonga	TACACTGGCA	TAGTCAACGA	GC6CGATTGG	GAAGACCGGG	GATAGGTTAT	TGCAATTATT	CGTCAACGAG	GAATGCCCAAG	TAAAGGCGAG	TCATCAGCTC	[1187]
16_Cryptoglana_pigra	TACATTGACA	CAGTGAAGGT	GTGTGGCTGG	GAATACTTGG	GATAGACCTT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCAGG	TCAT-AGCCT	[1150]
17_Cryptoglana_skujai	TACATTGACA	CAGTGAAGGT	GTGTGGCTGG	GAATACTTGG	GATAGACCTT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCAGG	TCAT-AGCCT	[1150]
18_Colacium_sp	TACATTGACA	CAGGGAAGGA	GTGTGAATGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCAAG	TCATCAGCTT	[1195]
19_Colacium_vesiculosum	TACATTGACA	CAGGGAAGGA	GTGTGAATGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCAAG	TCAT-AGCTT	[1118]
20_Euglena_clara	TACACTGGCC	CGGTGAAGGA	GTCCCCCTAT	ATNACSSSSS	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGTGCCGCG	-CTGCGGTGT	CCCTCCTCTG	[1192]
21_Discoplastis_adunca	TACATTGACA	CAGTGAATGT	GC6TGGATGG	GCATACTGGG	GATAGACGTT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCAGG	TCAT-AGCCT	[1191]
22_Euglena_deses	TACATTGATT	CAGTGAAGGT	GTGTGGGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGTC	GAATGCCTAG	TAGGCGTGAG	TCATCAGCTC	[1184]
23_Euglena_geniculata	TACATTGACA	CAGTGAATGA	GTGTGGTCTG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGCG	GAATACCTAG	TATGCGCAGG	TCATCAGCCT	[1200]
24_Euglena_granulata	TACATTGGTG	CAGTGAAGGA	GTGTGGCTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGTC	GAATGCCTAG	TATGCTTGG	TCATCAGCCA	[1188]
25_Euglena_hiemalis	TACATTGTCA	CAGTGAAGGT	GTGTGGATGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGTC	GAATGCCTAG	TATGCTTGG	TCATCAGCCC	[1184]
26_Eutreptia_pertyi	TACAATGATA	CAATGAAAGC	GC6TGGATGG	GCATGCTCGG	GATCGATTGCT	TGCAACTGTT	CGTTAACGAG	GAATGCCTAG	TATACGCACG	TCAT-ACCTT	[1160]
27_Euglena_stellata	TACACTGATC	GAGTGAAGGA	GTGTGGCTGG	GAATGCTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGTC	GAATGCCTAG	TAGGCGTTGG	TCACCAAGCA	[1165]
28_Euglena_longa	TACATTGTCA	CAGTGAAGGT	GTGAGGATGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGTC	GAATGCCTAG	TATGCTTGG	TCATCAGCCC	[1149]
29_Euglena_viridis	TACATTGACT	CAGTGAAGGA	GTGTGGCTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGTC	GAATGCCTAG	TAGGCGTTGG	TCATCAGCCA	[1196]
30_Euglena_exilis	TACATTGACT	CAGTGAAGGT	GTGGGGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATACCTAG	TATGCTTGG	TCATCAGCCC	[1195]
31_Euglena_gracilis	TACATTGTCA	CAGTGAAGGT	GTGTGGATGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGTC	GAATGCCTAG	TATGCTTGG	TCATCAGCCC	[1149]
32_Lepocinclis_steinii	TACATTGGCC	CAATGAATAT	GC6TGGGTGG	GGATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TAAGTGTGGG	TCATCAGCCC	[1153]
33_Lepocinclis_acus_var_major	TACATTGGTC	CAGTGAATGT	GTGTGGGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TAAAGCTGGG	TCAT-AGCCC	[1191]
34_Lepocinclis_tripteris	TACATTGATC	TGGTGAATGT	GTGTGGGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCAGG	TCAT-AGCCT	[1109]
35_Lepocinclis_spirogyroides	TACATTGGCC	CAGTGAATGT	GC6TGGGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATACCTAG	TAAAGCGCAGG	TCAT-AGCCT	[1110]
36_Lepocinclis_ovum	TACATTGACC	CAGTGAATAT	GTGTGAATGG	GGATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TAGGCGTGGG	TCAA-AGCCC	[1188]
37_Monomorphina_aenigmatica	TACATTGATA	CAGTGAAGGT	GTGTGGGTGG	GAATCCTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TAGGCGTGCG	TCATCAGCAC	[1188]
38_Colacium_mucronatum	TACATTGACA	CAGGGAAGGA	GTGTGAGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCAAG	TCATCAGCTT	[1174]
39_Monomorphina_sp	TACATTGACA	CAGCCAGGGT	GTGAGGGTGA	GAATCCTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTTGG	TAAAGCTGTG	TCAT-AGGCG	[1182]
40_Peranema_trichophorum	TACACTGGCA	AGGACACTTA	GC6TTAACCC	TCATGTCAGG	GATAGATTAT	TGCAATTATC	CACCAACGAG	GAATACCTAG	TAAAGCATAGG	TCATCGGCCCT	[1187]
41_Phacus_orbicularis	TACATTGGTC	TGGTGAATGT	GTGAGGGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TAGACGTGGG	TCATCAGCCC	[1161]
42_Phacus_similis	TACATTGGTC	CTGTGAATGT	GTGTGGGTGG	GAATACTGGG	GATAGCTGAT	TGCAACTGTC	TGCCAACGTC	GAATGCCTAG	TATGCTTGGG	TCAT-AGCCC	[1157]
43_Monomorphina_pyrum	TACATTGACA	CAGCCAGGGT	G-GAGGGTGA	GAATCCTGGG	GATAGACTCT	TGCAACTGTC	TGCCAACGAG	GAATGCCTTGG	TAAAGGTGTG	TCAT-AGTGC	[1178]
44_Strombomonas_acuminata	TACATTGACA	CAGTGAATGT	GTGTGGCTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGTGAG	TCAT-AGCTC	[1119]
45_Strombomonas_napiformis	TACATTGACA	CAGTGAATGT	GTGTGGCTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGTGAG	TCAT-AGCTC	[1198]
46_Strombomonas_borysteniensis	TACATTGACA	CAGTGAAGGT	GTGTGGGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGTGAG	TCAT-AGCTC	[1196]
47_Strombomonas_ovalis	TACATTGACA	CAGTGAAGGT	GTGTGGGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGTGAG	TCAT-AGCTC	[1196]
48_Trachelomonas_echinata	TACATTGACA	CAGTGAAGGT	GC6TGAAGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCTTGTG	TCATCAGCAC	[1198]
49_Strombomonas_triquetra	TACATTGACA	CAGTGAAGGT	GTGTGTGTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGTGAG	TCAT-AGCTC	[1199]
50_Eutreptiella_pomquetensis	TACACTGACA	CAGTGAACGA	GC6TGGGCGG	GAATGCTGGG	GATAGATTAT	TGCAACTATT	TGCCAACGAG	GAATGCCTAG	TAAAGCGCAAG	TCATCAGCTT	[1163]
51_Trachelomonas_lefevrei	TACATTGACA	CAGTGAAGGT	GTGTGAATGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TAGGCGTGGG	TCATCAGCCC	[1187]
52_Trachelomonas_planctonica	TACATTGACA	CAGTGAAGGT	GTGTGATTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TAGGCGTGGG	TCATCAGCCC	[1194]
53_Trachelomonas_similis	TACATTGATT	CAGTGAAGGT	GTGTGGGTGG	GAACGATGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGCGTG	TCAT-AGCAC	[1186]
54_Trachelomonas_volvocina	TACATTGATT	CAGTGAAGGT	GTGTGAATGG	GAACGCTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TATGCGGTG	TCAT-AGCAC	[1162]
55_Trachelomonas_conradi	TACATTGACA	CAGTGAAGGT	GTGTGGCTGG	GAATACTGGG	GATAGATTAT	TGCAACTGTC	TGCCAACGAG	GAATGCCTAG	TAGGCGTGGG	TCATCAGCCC	[1191]

	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	
01_Rhynchomonas nasuta	GtGcCgAtTa	CgTcCcTgCc	AtTtGtAcAc	AcCgCcCgTc	GtTgTtTcCg	AtGaTgGtGc	TgCaGgTgAt	CgGaCaGtTc	AcCgAcAtTt	TtCaAtAgAg	[1290]
02_Blastocrithidia culicis	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	ATTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACGGTTC	ACCGATATTT	TTCAATTAGAG	[1279]
03_Herpetomonas muscarum	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGATATTT	TTCAATTAGAG	[1276]
04_Leishmania braziliensis	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGATATTT	TTCAATA---	[1280]
05_Leptomonas bifurcata	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGATATTT	TTCAATTAGAG	[1277]
06_Leishmania infantum	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGAT---	-----	[1258]
07_Leishmania major	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGATATTT	TTCAATTAGAG	[1283]
08_Leishmania tarentolae	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGATATTT	TTCAATTAGAG	[1289]
09_Trypanosoma cruzi	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGATATTT	TTCAATTAGAG	[1282]
10_Trypanosoma brucei	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGATATTT	TTCAATTAGAG	[1293]
11_Bodo saltans	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGGTGAT	CGGACAGTTC	ACCGATATTT	TTCAATTAGAG	[1287]
12_Dimastigella trypaniformis	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TGCAGGTGAT	CGGACAGTTC	ACCGACATTT	TTCAATTAGA-	[1284]
13_Rhynchobodo sp	GTGCCGATTA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TACAGATGAT	CGGACTGTTC	ATCGATATTT	TTCAATTAGAG	[1266]
14_Colacium elongatum	GCACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGGCAG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCCACTAGAG	[1298]
15_Ploetia oblonga	GCCCTGATTG	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGTTTCCG	ATGATGGTGG	TGCAGGTGAT	GAGAGGCTTC	TCCAATATGC	ACAGCTAGAG	[1287]
16_Cryptoglena pigra	GCACAGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCAGCAG	TAGAGTCATC	AGGAGGATTG	GACAATGTCC	GCTGCTAGAG	[1250]
17_Cryptoglena skujai	GGACAGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCAGCAG	TAGAGTCATC	AGGAGGATTG	GACAATGTCC	GCTGCTAGAG	[1250]
18_Colacium sp	GCACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGGCAG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCCACTAG-A	[1294]
19_Colacium vesiculosum	GCACCGATTG	TGTCCTTGCC	-----	-----	-----	-----	-----	-----	-----	-----	[1138]
20_Euglena clara	TGTC--CTG	CCCTGAAGAG	GGTGGGAAAT	CCTGCAAGTC	CGTTTGTGAG	AGCCGTGCTG	TAGATGATTG	CAACTGCACC	GATTGTGTCC	GCCATTTGTA	[1290]
21_Discoplastis adunca	GCACCGATTG	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCAGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTGCTANNNG	[1291]
22_Euglena deses	ACACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCAGCTG	TAGAGCCATC	CGGAGGATTN	GGTGTGTTT	GCTGCTAGAG	[1284]
23_Euglena geniculata	GCACAGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGTGTT	TAGAGCCATC	AGGAGGATTG	GGTAATGTCC	GCTGCTAG-A	[1299]
24_Euglena granulata	AGACCGATTG	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGACAGCTG	TAGAACCATC	AGGAGGATTG	GGTGTATGCC	GCTGTTAGAG	[1288]
25_Euglena hiemalis	AGACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGGCTG	TAGAGTCATC	AGGAGGATTG	GACAATGTCC	GCTGCTAGAG	[1284]
26_Eutreptia pertyi	GCATCGACTG	CGTCCCTGCC	AATTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGTGTT	CAGAGACTTC	AGGAGATTGG	GCCATGTGTC	GCGGCTAGAG	[1260]
27_Euglena stellata	ACACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGATCGCTG	TAGAACCATC	AGGAGGATTG	GGTGTATGCC	GTGACTAGAG	[1265]
28_Euglena longa	AGACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGGCTG	TAGAGTCATC	AGGAGGATTG	GACAATGTCC	GCTGCTAGAG	[1249]
29_Euglena viridis	ACACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGATAGCTG	TAGAACCATC	AGGAGGATTG	GGTAATGTCC	GCTGTTAG-A	[1295]
30_Euglena exilis	ACACAGACTG	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCATCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTGCTAGAG	[1295]
31_Euglena gracilis	AGACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGGCTG	TAGAGTCATC	A-----	-----	-----	[1220]
32_Lepocinclis steinii	GCACTGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGTGTT	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTACTAGAG	[1253]
33_Lepocinclis acus var major	ACACTGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGGCAG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCCACTAGAG	[1291]
34_Lepocinclis tripteris	GCACCGATTG	TGTCCTTGCC	-----	-----	-----	-----	-----	-----	-----	-----	[1129]
35_Lepocinclis spirogyroides	GCACTGATTG	TGTCCTTGCC	-----	-----	-----	-----	-----	-----	-----	-----	[1130]
36_Lepocinclis ovum	ACGTGCTATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCAGCTG	TAGAGCCATC	AGGAGGATTG	GGCGATGTCC	GCTGCTAGAG	[1288]
37_Monomorphina aenigmatica	ACACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGAAAGCAG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTTCTAGAG	[1288]
38_Colacium mucronatum	GCACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGGCAG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCCACTAGAG	[1274]
39_Monomorphina sp	ACCCTGAATA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGTGTT	CAGAGCCATC	CGGAGGATTG	GGCAATGTCA	GTCACTAGAG	[1282]
40_Peranema trichophorum	ATATTGATTA	CGTCCCTGCC	ACTTGTACAC	ACCGCCCGTC	GTTGCTACTG	ATGGTATGTT	CAGAGGCTGG	AAGAGATCCA	TCCAATGTGG	CATGCTAGAA	[1287]
41_Phacus orbicularis	GCATCGATTG	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGTGTT	TAGAGCCATC	CGGAGACTCG	GGCAATGTCT	GCTACTAGAG	[1261]
42_Phacus similis	AAACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCAGCAG	TAGAGCCATC	AGGAGGATTG	GGCGATGTCC	GCTGCTAGAG	[1257]
43_Monomorphina pyrum	ACCCTGAATA	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGTTGATGG	CAGAGCCATC	CGGAGGATTG	GGCAATGTCA	GTCAATAGAG	[1278]
44_Strombomonas acuminata	ACACAGATTG	CGTCCCTGCC	-----	-----	-----	-----	-----	-----	-----	-----	[1139]
45_Strombomonas napiformis	ACACAGATTG	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ACGGCAGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GTTGCTGAAAG	[1298]
46_Strombomonas borysteniensis	ACACAGATTG	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCAGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTGCTAAAAG	[1296]
47_Strombomonas ovalis	ACACAGATTG	CGTCCCTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGCAGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTGCTAAAAG	[1296]
48_Trachelomonas echinata	AAACAGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGAGAGCAA	CAGAGTCATC	AGGAGGATTG	GGCAATGTCC	GCTGCTAGAG	[1298]
49_Strombomonas triquetra	ACACAGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGGTGTGTT	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTGCTAAAAG	[1299]
50_Eutreptiella pomquetensis	GCACTGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGATAGAAG	TAGAGGCATC	AGGAGGTTTG	GGCAATGTCC	GCTATTAGAG	[1263]
51_Trachelomonas lefevrei	ACACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGATCGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCGATTAGAG	[1287]
52_Trachelomonas planctonica	ACACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGATCGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GTGATTAGAG	[1294]
53_Trachelomonas similis	GCACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGACAGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTGTTAGAG	[1286]
54_Trachelomonas volvocina	GCACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGACAGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTGTTAGAG	[1262]
55_Trachelomonas conradi	ACACCGATTG	TGTCCTTGCC	ATTTGTACAC	ACCGCCCGTC	GTTGCTACCG	ATGACAGCTG	TAGAGCCATC	AGGAGGATTG	GGCAATGTCC	GCTGTTAGAG	[1291]

	1310	1320	
01_Rhynchomonas_nasuta	GaAgCaAaAg	TcGtAaCaAg	GtAg [1314]
02_Blastocrithidia_culicis	GAAGCAAAAG	TCGT-----	---- [1293]
03_Herpetomonas_muscarum	GAAGCAAAAG	T-----	---- [1287]
04_Leishmania_braziliensis	-----	-----	---- [1280]
05_Leptomonas_bifurcata	GAAGCAAAAG	TC-----	---- [1289]
06_Leishmania_infantum	-----	-----	---- [1258]
07_Leishmania_major	GAAGCAA---	-----	---- [1290]
08_Leishmania_tarentolae	GAAGCAAAAG	TCGTAACAAG	GTAG [1313]
09_Trypanosoma_cruzi	GAAGCAAAAG	-----	---- [1292]
10_Trypanosoma_brucei	GAAGCAAAAG	TCGTAACAAG	GTAG [1317]
11_Bodo_saltans	GAAGCAAAAG	TCGTAACAAG	GTAG [1311]
12_Dimastigella_trypaniformis	-----	-----	---- [1284]
13_Rhynchobodo_sp	GAAGCAAA--	-----	---- [1274]
14_Colacium_elongatum	GAAGCAAAAG	TCGTAACAAG	GTTG [1322]
15_Ploeotia_oblonga	ACGATAAAAG	TCGTAACAAG	GTAG [1311]
16_Cryptoglana_pigra	GAAGCAAAA-	-----	---- [1259]
17_Cryptoglana_skujai	GAAGCAAAA-	-----	---- [1259]
18_Colacium_sp	GAAGCAAAAG	TCGTAACAAG	GTTG [1318]
19_Colacium_vesiculosum	-----	-----	---- [1138]
20_Euglena_clara	CCACCGCCCG	TCGTTGCTAC	GATG [1314]
21_Discoplastis_adunca	GAAGCAAAA-	-----	---- [1300]
22_Euglena_deses	GAAGCAAAA-	-----	---- [1293]
23_Euglena_geniculata	GAAGCAAAAG	TCGTAACAAG	GTTG [1323]
24_Euglena_granulata	GAAGCAAAA-	-----	---- [1297]
25_Euglena_hiemalis	GAAGCAAAAG	TCGTAACAAG	GTTG [1308]
26_Eutreptia_pertyi	GAAGCAAAAG	TCGTAACAAG	GCTG [1284]
27_Euglena_stellata	GAAGCAAAAG	TCGTAACAAG	GTTG [1289]
28_Euglena_longa	GAAGCAAAA-	-----	---- [1258]
29_Euglena_viridis	GAAGCAAAAG	TCGTAACAAG	GTTG [1319]
30_Euglena_exilis	GAAGCAAAAG	TCGTAACAAG	GTTG [1319]
31_Euglena_gracilis	-----	-----	---- [1220]
32_Lepocinclis_steinii	GAAGCAAAAG	TCGTAACAAG	GTTG [1277]
33_Lepocinclis_acus_var_major	GAAGCAAAAG	TCGTAACAAG	GTT- [1314]
34_Lepocinclis_tripteris	-----	-----	---- [1129]
35_Lepocinclis_spirogyroides	-----	-----	---- [1130]
36_Lepocinclis_ovum	GAAGCAAAAG	TCGTAACAAG	GTTG [1312]
37_Monomorphina_aenigmatica	GAAGCAAAAG	TCGTAACAAG	G-TG [1311]
38_Colacium_mucronatum	GAAGCAAAAG	TCGTAACAAG	GTTG [1298]
39_Monomorphina_sp	GAAGCAAAAG	TCGTAACAAG	GTCG [1306]
40_Peranema_trichophorum	GAAGCAAAAG	TCGTAACAAG	GTTG [1311]
41_Phacus_orbicularis	GAAGCAAAA-	-----	---- [1270]
42_Phacus_similis	GAAGCAAAA-	-----	---- [1266]
43_Monomorphina_pyrum	GAAGCAAAAG	TCGTAACAAG	GT-- [1300]
44_Strombomonas_acuminata	-----	-----	---- [1139]
45_Strombomonas_napiformis	GAAGCAAAAG	TCGTAACAAG	GTTG [1322]
46_Strombomonas_borysteniensis	GAAGCAAAAG	TCGTAACAAG	GTTG [1320]
47_Strombomonas_ovalis	GAAGCAAAAG	TCGTAACAAG	GTTG [1320]
48_Trachelomonas_echinata	GAAGCAAAAG	TCGTAACAAG	GTTG [1322]
49_Strombomonas_triquetra	GAAGCAAAAG	TCGTAACAAG	GTTG [1323]
50_Eutreptiella_pomquetensis	GAAGCAAAA-	-----	---- [1272]
51_Trachelomonas_lefevrei	GAAGCAAAAG	TCGTAACAAG	GTTG [1311]
52_Trachelomonas_planctonica	GAAGCAAAAG	TCGTAACAAG	GTTG [1318]
53_Trachelomonas_similis	GAAGCAAAAG	TCGTAACAAG	GTTG [1310]
54_Trachelomonas_volvocina	GAAGCAAAA-	-----	---- [1271]
55_Trachelomonas_conradi	GAAGCAAAAG	TCGTAACAAG	GTTG [1315]

Figure 3.2. The 28S interleaved alignment. Kinetoplastids are listed first and then euglenids, with an identifying number for each OTU to the left of the species name. The column positions, from 1-1452, are given in ten base increments at the top of each page. The number of residues from the original sequence is given in brackets to the right of each line corresponding to its respective OTU.

	10	20	30	40	50	60	70	80	90	100	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
05_Leishmania_infantum	GCATATTACT	CAGCGGAGGA	AAAGAAAACA	ACCGTGATTG	TCTCAGTAGC	GGCGAGCGAA	GAGGGAAAA-	-TC--GAG--	TGAGTATTGA	GCCAAAGAAG	[94]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
07_Leishmania_major	GCATATTACT	CAGCGGAGGA	AAAGAAAACA	ACCGTGATTG	TCTCAGTAGC	GGCGAGCGAA	GAGGGAAAA-	-TC--GAG--	TGAGTATTGA	GCCAAAGAAG	[94]
08_Leishmania_tarentolae	GCATATTACT	CAGCGGAGGA	AAAGAAAACA	ACCGTGATTG	TCTCAGTAGC	GGCGAGCGAA	GAGGGAAAA-	-TC--GAG--	TGAGTATTGA	GCCAAAGAAG	[94]
09_Trypanosoma_cruzi	GCATATTACT	AAGCGGAGGA	AAAGAAAACA	ACCGTGATTG	TCTCAGTAGC	GGCGAGCGAA	GAGGGATAA-	-TC--GAA--	TGAGTATTGA	GCCAAAGAAG	[94]
10_Trypanosoma_brucei	GCATATTACT	CAGCGGAGGA	AAAGAAAACA	ACCGTGATTG	TCTCAGTAGC	GGCGAGCGAA	GAGGGATAA-	-TC--GCG--	TGAGTATTGA	GCCAAAGAAG	[94]
11_Bodo_saltans	-----	-----	-----	-----	-----	GGG	CCTCAAAA-	AGAGTATTGA	GCCAAAGAAG	[31]	
12_Dimastigella_trypaniformis	GCATATTATT	CAGCGGAGGA	AAAGAAAACA	ACCGTGATTG	CCTCAGYAGC	GGCGAGCGAA	GAGGGATAA-	-TC--GATGG	TGAGTATTGA	GCCATAGAAG	[96]
13_Rhynchobodo_sp	GCATATTATT	CAGCGGAGGA	AAAGAAAACA	ACCGTGATTG	CCTCAGTAGC	GGCGAGCGAA	GAGGGATAA-	-TC--GATGG	TGAGTATTGA	GCCATAGAAG	[96]
14_Ploeotia_oblonga	GCATATCACT	CAGTGGAGGA	AAAGAAAGCA	ACTGCGATTG	TCTCAGTAAC	GGCGAGTGAA	CAGAGCAAG-	-TG--AGC--	CGAACGGCAC	ACCGCAGAGG	[94]
15_Colacium_elongatum	GCATATCACT	CAGTGGAGGA	AGAGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	CTGGCAACGC	AGACGCGCAG	GGAATGGTCC	ACGAGAGAAG	[100]
16_Cryptoglena_pigra	GCATATCACT	CAGTGGAGGA	AAGGCAAAACA	ACTGTGATTG	CCCCAGTAAT	GGCGAGTGAA	CAGGCAATGG	CTAT-CTCCA	TGAATGGCTC	ACCCAGAGG	[99]
17_Cryptoglena_sukjai	-----	-----NAGGA	AAGGCAAAACA	ACTGTGATTG	CCCCAGTAAT	GGCGAATGAA	CAGGCAATGG	CTATCCTCCA	GGAATGGCTC	ACCCAGAGG	[85]
18_Colacium_sp	GCATATCACT	CAGTGGAGGA	AGAGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	CTGGCAAAGT	GATTCCGCGG	GGAATGGTCC	ACGAGAGAAG	[100]
19_Colacium_vesiculosum	GCATATCACT	CAGTGGAGGA	AGAGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	CTGGCAAAGT	GTACACGCGA	GGAATGGTCC	ACGAGAGAAG	[100]
20_Euglena_clara	AAACCGAAGT	GCGTGGTGAA	GGCAACACCA	AGGGAGAGAG	GGGGAGGAGG	GGGGA--GGAG	GAGGTAATGG	-TG--CGGGA	GGTCGTCTGCT	GTGCGCGTCT	[96]
21_Discoploastis_adunca	GCATATCACT	CAGTGGAGGA	ACAGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	CTGGCAGAG-	-ATTCCGCGA	GGAATGGCTC	ACGACAGAGG	[98]
22_Euglena_deses	GCATATCACT	CAGTGGAGGA	AATGCAAGCA	ACCGCGATTG	CTGGAGTAAC	GGCGAGTGAA	CGGGCAGTGG	GAATTCGG--	GGAATGGGCC	ACGACAGAGG	[98]
23_Euglena_geniculata	GCATATCACT	CAGTGGAGGA	ACAGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAATGAA	CAGGCAAGAG	TTCCCTG-CG	GGAACGGGCA	GCCATAGAAG	[99]
24_Euglena_granulata	GCATATCACT	CAGTGGAGGA	AATGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	CGAGGAAGAT	TATCTCGCGG	GGAAGGGGTC	ACCACCGAGG	[100]
25_Euglena_hiemalis	GCATATCACT	CAGTGGAGGA	ACAGCAAGCA	ACTGCGATTG	CTGCAGTAAT	GGCGAATGAA	CGAGCAGAT-	-CGGGAGCGA	CGAATGACCC	ACAGAACAGG	[98]
26_Eutreptia_pertyi	GCATATCACT	CAGTGGAGGA	TGAGAAAGTA	ACAACGATTG	CGGCAGTAAT	GGCGAATGAA	GGTGCACAGA	CGGTGCC-CG	GGAATGGCGC	GCGGTACAGG	[99]
27_Euglena_stellata	-----	-----NAGGA	AAGGAAAGCA	ACTGCGATTG	CTCCAGKAAT	GGCGAATGAA	ATGGCATAGT	GAGTTTGCCG	TGAATGGGCT	ACCACAGAGG	[85]
28_Euglena_longa	-----	-----NAGGA	ACAGCAAGCA	ACATCGATTG	CTGCAGTAAT	GGCGAATGAA	CGAGCAGTGG	GACTCCGCGG	TGAATGACCC	ACAGAACAGG	[85]
29_Euglena_viridis	GCATATCACT	CAGTGGAGGA	AATGCAAGCA	ACAGCGATTG	CTTCAGTAAT	GGCGAATGAA	CCAGCAAAAG	AAGGTTGA--	AGAAAGGGCC	ACAATAGAAG	[98]
30_Euglena_exilis	GCATATCACT	CAGTGGAGGA	AGCGCAAGCA	ACGCGCGATTG	CCGCAGTAAC	GGCGAGTGAA	ATGGCAACTG	GAGATCGGCG	GGAACGGCTC	ACTACACAGG	[100]
31_Euglena_gracilis	GCATATCACT	CAGTGGAGGA	ACAGCAAGCA	ACCTCGATTG	CTGCAGTAAT	GGCGAATGAA	CGAGCAGGGG	CTTCTCGCAG	TGAATGACCC	ACGAAACAGG	[100]
32_Lepocinclis_steinii	-----	-----NAGGA	AGCCCAAGCA	ACTGCGATTG	CTTCAGTAAT	GGCGATTGAA	CTGGCATAA-	-AGAGAGT--	GGAATGGTTT	GCAACAGAGG	[81]
33_Lepocinclis_acus_var_major	-----	-----NAGGA	AGCCCAAGCA	ACTGCGATTG	CTTCAGTAAT	GGCGATTGAA	CTGGCATCG-	-AGATTGTCTG	GGAATGGTCC	ACACGAGAGG	[83]
34_Lepocinclis_spirogyroides	-----	-----NAGGA	AGACCAAGCA	ACCGCGATTG	CTTCAGTAAC	GGCGAGTGAA	CCGGCACCGA	GAT--TGCGG	TGAAGGGGCC	ACAACAGAGG	[83]
35_Lepocinclis_triapteris	-----	-----NAGGA	ATCCCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	CTGGCATAGA	TTCCCTGCA	GGAATGGTCC	ACAACAGAGG	[85]
36_Lepocinclis_ovum	-----	-----NAGGA	AGCCCAAGCA	ACTGCGATTG	CTACAGTAAT	GGCGATTGAA	CTGGCATGGA	GATTGAGTGG	GGAATGGTCC	ACAACAGAGG	[85]
37_Monomorphina_aenigmatica	GCATATCACT	CAGTGGAGGA	AAGGCAAGCA	ACGGCGACTG	CCTCAGTAAC	GGCGAGTGAA	TTGGCAAAGT	GGAACCGCAG	GGAATGGCCC	ACAGCAGAGG	[100]
38_Colacium_mucronatum	GCATATCACT	CAGTGGAGGA	AGAGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	CTGGCAAAGT	GATTCCGCGA	GGAATGGTCC	ACGAGAGAAG	[100]
39_Monomorphina_sp	GCATATCACT	CAGTGGAGGA	AAGACAGGCA	ACCGCGATTG	CCCCAGTAAT	GGCGAATGAA	TTGGCAAAGT	GGTTTCCTG-	GGAATGGCCC	ACATCAGAGG	[99]
40_Peranema_trichophorum	GCATATCACT	CAGTGGAGGA	AAGGAAAGAA	ACATCGACGG	TCTTAGTAAA	GGCGATTGAA	GAGGCCAAC-	-CA--CGA--	CGAAAGGCC	ATCCTCAGGG	[94]
41_Phacus_orbicularis	GCATATCACT	CAGTGGAGGA	ATGCCAAGCA	ACCGCGATTG	CTCTAGTAAC	GGCGAGTGAA	CCGGCAAC-G	GATCCGAGGA	GGAATGGACC	GCCACGAGG	[99]
42_Phacus_similis	-----	-----NAGGA	AGCCCAAGCA	ACTGCGATTG	CTACAGTAAT	GGCGAGTGAA	GTGGCATAG-	-AT--CGT--	GGAATGGCAT	ACAACAGAGG	[79]
43_Monomorphina_pyrum	-----	-----NAGGA	AAGCAAGCA	ACTGCGATTG	CCCCAGTAAT	GGCGAATGAA	TTGGCAAAGT	GTTCCTCGCG	GGAATGGCCC	ACGACAGAGG	[85]
44_Strombomonas_napiformis	GCATATCACT	CAGTGGAGGA	AGTGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	TGGGCAAAA-	-TC--CGCAG	GGAAGGGCCA	ACCAGAGCGG	[96]
45_Strombomonas_borysteniensis	GCATATCACT	CAGTGGAGGA	AATGCAAGCA	ACCGCGATTG	CTTCAGTAAC	GGCGAGTGAA	TGGGCAAAAT	-CC--CGCAG	GGAATGGCCT	ACCACAGTGG	[97]
46_Strombomonas_acuminata	-----	-----NAGGA	AGTGCAAGCA	ACCGCGATTG	CTTCAGTAAC	GGCGAGTGAA	TGGGCAAAAT	-CC--CGCAG	GGAAGGGCCT	ACCAGAGTGG	[82]
47_Strombomonas_ovalis	GCATATCACT	CAGTGGAGGA	AATGCAAGCA	ACCGCGATTG	CTTCAGTAAC	GGCGAGTGAA	TGGGCAAAAT	-CC--CGCAG	GGAATGGCCT	ACCACAGTGG	[97]
48_Trachelomonas_echinata	-----	-----NAGGA	AAGCCAAGCA	MCAGCGATTG	CTTCAGTAAT	GGCGAATGAA	CTGGCAAAA-	-TC--CGCA	GGAATGGGCT	ACTACAGAAG	[81]
49_Strombomonas_triquetra	GCATATCACT	CAGTGGAGGA	ACGCAAGCA	ACCGCGATTG	CTTCAGTAAC	GGCGAGTGAA	TGGGCAAGG-	-AT--CGCAG	GGAAGGGCCT	ACCAGAGTGG	[96]
50_Eutreptiella_pomquetensis	GCATATCACT	CAGTGGAGGA	GAAGAAAGTA	ACTACGATGG	CCCCAGTAAT	GGCGAATGAA	CAGGCACAGG	GGGCTTGACG	GGAACGGCGC	GCAGCAAAAG	[100]
51_Trachelomonas_planctonica	GCATATCACT	CAGTGGAGGA	AATGCAAGCA	ACAGCGATTG	CTGCAGTAAC	GGCGAGTGAA	TGGGCAAGT	AAT--CGT--	GGAATGGTCC	ACGTTACAGG	[96]
52_Trachelomonas_lefevrei	GCATATCACT	CAGTGGAGGA	AATGCAAGCA	ACAGCGATTG	CTGCAGTAAC	GGCGAGTGAA	TGGGCAAGT	GAATCCGT--	GGAATGGTCC	ACGATACAGG	[98]
53_Trachelomonas_volvocina	GCATATCACT	CAGTGGAGGA	AGCGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	TGGGCAAAGC	GAGTTCGCGA	GGAAGGCTC	ACCGCAGAGG	[100]
54_Trachelomonas_conradi	-----	-----NAGGA	AGAGCAAGCA	ACTGCGATTG	CCACAGTAAC	GGCGAGTGAA	TGGGCAAAGT	GAAT-CTC--	GGAATGGTCC	ACGACACAGG	[82]
55_Trachelomonas_similis	GCATATCACT	CAGTGGAGGA	AAGCAAGCA	ACTGCGATTG	CTTCAGTAAC	GGCGAGTGAA	TTGGCAAAGT	GAAACGGCAG	AGAAGGGCCC	ACAAGAGAGG	[100]

	110	120	130	140	150	160	170	180	190	200	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-Ga	AgAcCgAtAg	CgAaCaAgTa	GcGtGcAaGg	[32]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-----	-CGA	AGACCGATAG	CGAACCAAGTA	GCCTGAGCGA	[33]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
05_Leishmania_infantum	GTGTTAGCCC	ATTGAGTACA	CTGTTTGGGA	ATGCAAGTGC	AAGAGGTATT	TGTGCCAAAG	TTAAATACGA	AGACCGATAG	CGAACCAAGTA	GCCTGAGCGA	[194]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-GA	AGACCGATAG	CGAACCAAGTA	[32]
07_Leishmania_major	GTGTTAGCCC	ATTGAGTACA	CTGTTTGGGA	ATGCAAGTGC	AAGAGGTATT	TGTGCCAAAG	TTAAATACGA	AGACCGATAG	CGAACCAAGTA	GCCTGAGCGA	[194]
08_Leishmania_tarentolae	GTGTTAGCCC	ATTGAGTACA	CTGTTTGGGA	ATGCAAGTGC	AAGAGGTATT	TGTGCCAAAG	TTAAATACGA	AGACCGATAG	CGAACCAAGTA	GCCTGAGCGA	[194]
09_Trypanosoma_cruzi	GTGTTAGCCC	ATTGAGTACA	CTGTTTGGGA	ATGCAAGTGC	AAGAGGTATT	TGTGCCAAAG	TTAAATACGA	AGACCGATAG	CGAACCAAGTA	GCCTGAGCGA	[194]
10_Trypanosoma_brucei	GTGTCAGCCC	ATTGAGTACA	CTGTTTGGGA	ATGCAAGTGC	AAGAGGTATT	TGTGCCAAAG	TTAAATATGA	AGACCGATAG	CGAACCAAGTA	GCCTGAGCGA	[194]
11_Bodo_saltans	GTGTTAGCCC	ATTGAGTACA	CTGTTTGGGA	ATGCAAGTGC	AAGAGGTATT	TGTGCCAAAG	TTAAATATGA	AGACCGATAG	CGAACCAAGTA	GCCTGAGCGA	[131]
12_Dimastigella_trypaniformis	GTGTCAGCCC	ATTGAGTACA	CTGTTTGGGA	ATGCAAGTGC	AAGAGGTATT	TGTGCCAAAG	TTAAATACGA	AGACCGATAG	CGAA-----	-----	[180]
13_Rhynchobodo_sp	GTGTCAGCCC	ATTGAGTACA	CTGTTTGGGA	ATGCAAGTGC	AATGGTATGT	GTTGCCAAAG	TTAAATACGA	AGACCGATAG	CGAA-----	-----	[180]
14_Ploeotia_oblonga	GTGATAGTCG	CGTGAGTCGG	CTGTTTGGGA	CTGCAGTCT	ACGGATGTAC	CCTCTAAGG	TTGAAGGCCA	AGACCGATAG	CGAACCAAGTA	GAGTGATCGA	[194]
15_Colacium_elongatum	GATACTGTCC	TGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	AAGCATACCG	AGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[200]
16_Cryptoglena_pigra	GATACTGTCC	CCCGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	CTAAACACTG	GAACCGACAG	AGCACAAAGTA	AAGCGATTGA	[199]
17_Cryptoglena_sujai	GATACTGTCC	CCTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	CTAAACTACTG	GAACCGACAG	AGCACAAAGTA	AAGCGATTGA	[185]
18_Colacium_sp	GATACTGTCC	TGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	AAGAATACCG	AGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[200]
19_Colacium_vesiculosum	GATACTGTCC	TGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	ACGAATACCG	AGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[200]
20_Euglena_clara	GTGCGCGCCC	YG-CTACCCC	CCCCGTCTGC	CGCATCCCG	AT-GACATGC	GC-TGCAGGG	ATCCACCGGG	ATCCACCGGG	ACAGGGTTTC	GAGTGCTTGC	[193]
21_Discoplastis_adunca	GCTACTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGTTGTAT	CCCCACAAGG	CATCATACTG	GGACCGATAG	AGCACAAAGTA	GAGTGATTGA	[198]
22_Euglena_deses	GACACGGTCC	CCTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGTTGTAT	CCTCGCAAGG	AAGCATACTG	GAACCGATAG	CGCACAAAGTA	GAGTGATCGA	[198]
23_Euglena_geniculata	CGGAAAGCCC	CCTGAGTAAG	CTGTTTGGGA	CTGCAGTTTG	AAGTTGTAT	CCTCGCAAGG	CAGAATACTG	AGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[199]
24_Euglena_granulata	GCTACCGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGCGTGTAT	CCCCGCAAGG	AAGCATACTG	GGACCGATAG	CGCACAAAGTA	GAGTGATCGA	[200]
25_Euglena_hiemalis	GATACTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGTTGTAT	CCTCGCAAGG	TGGCATACTG	GGACCGATAG	CGCAGAAAGTA	GAGTGATCGA	[198]
26_Eutryptia_pertyi	GTGAAAGCCC	CGTGAGTAAG	CTGTTTGGGA	CTGCAGTTTG	AATTGTGTAT	CCTCGCAAGG	CTGAATATTG	GAACCGACAG	TGGACAAGTA	GAGTGATCGA	[199]
27_Euglena_stellata	GCTATCGTCC	CCTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	AGGAATACTG	GGACCGATAG	AGCACAAAGTA	GAGCGATCGA	[185]
28_Euglena_longa	GAGACTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGTTGTAT	CCTCGCAAGG	AGGCATACTG	GGACCGATAG	CGCAGAAAGTA	GAGTGATCGA	[185]
29_Euglena_viridis	GAAACTGTCC	CCTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	AAGCATACTG	AGACCGATAG	TGCACAAGTA	GAGTGATCGA	[198]
30_Euglena_exilis	GATACTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AACGGTGTAT	CCCCGCAAGG	AAGAACACCG	GGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[200]
31_Euglena_gracilis	GAGACTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGCGTGTAT	CCTCGCAAGG	TGGCATACTG	GGACCGATAG	TGCACAAGTA	GAGTGATCGA	[200]
32_Lepocinclis_steinii	GATATTGTCC	CGTGAGTAAG	CGGCTTGGGA	CTGCAGTTTG	AAGCGTGTAT	CCTCGCAAGG	TGGAATACTG	GGACCGATAG	AGCACAAAGTA	GAGTGATTGA	[181]
33_Lepocinclis_acus_var_major	GATATTGTCC	CGTGAGTAAG	CGGCTTGGGA	CTGCAGTTTG	AAGCGTGTAT	CCTCGCAAGG	TGGAATACTG	GGACCGATAG	CGCAGAAAGTA	GAGTGATTGA	[183]
34_Lepocinclis_spirogyroides	GATACTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGCGTGTAT	CCTCGCAAGG	TGGAATACTG	GGACCGACAG	AGCACAAAGTA	GAGTGATCGA	[183]
35_Lepocinclis_triptyris	GATATTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCCCGCAAGG	AGGAATACTG	GAACCGTATG	CGCACAAAGTA	GAGTGATTGA	[185]
36_Lepocinclis_ovum	GATATTGTCC	TGTGAGTAAG	CGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	CGAATACTG	GGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[185]
37_Monomorphina_aenigmatica	GATACTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	CAGAGTACTG	GGACCGATAG	AGCACAAAGTA	GAGTGATTGA	[200]
38_Colacium_mucronatum	GATACTGTCC	TGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCACGCAAGG	ACGAATACCG	AGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[200]
39_Monomorphina_sp	GAGAGTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	CAGAATACTG	GGACCGATAG	AGCACAAAGTA	GAGTGATCGA	[199]
40_Peranema_trichophorum	GTATAGTCC	CGTGAGTAAG	-AAGTTGAGC	GTGTT-TTG	AAGGGTGAAT	CCTCGCAAGG	CTAAGTACCG	AGACCGATAG	TAAACAAGTA	GAGCGATTGA	[192]
41_Phacus_orbicularis	GATACTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGCATGTAT	CCTCGCAAGG	TGCAATACTG	GGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[199]
42_Phacus_similis	GATACTGTCC	CCTGAGTAAG	CGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCACGCAAGG	AGGAATACTG	GGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[179]
43_Monomorphina_pyrum	GAGAGTGTCC	CATGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	CAGAATACTG	GGACCGATAG	AGCACAAAGTA	GAGTGATTGA	[185]
44_Strombomonas_napiformis	GATACTGTCC	TTTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTATAT	CCCCGCAAGG	CAGCATACTG	GGACCGATAG	TGCACAAGTA	GAGTGATCGA	[196]
45_Strombomonas_borsteniensis	GATACTGTCC	TTTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTATAT	CCTCGCAAGG	CAGCATACTG	GAACCGATAG	CGCACAAAGTA	GAGTGATCGA	[197]
46_Strombomonas_triquetra	GATACTGTCC	TTTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTATAT	CCCCGCAAGG	CAGCATACTG	GGACCGATAG	TGCACAAGTA	GAGTGATCGA	[182]
47_Strombomonas_ovalis	GATACTGTCC	TTTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTATAT	CCTCGCAAGG	CAGCATACTG	GAACCGATAG	TGCACAAGTA	GAGTGATCGA	[197]
48_Trachelomonas_echinata	GATACTGTCC	TGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCCTGCAAGG	CAGAATACTG	GAACCGATAG	CGCACAAAGTA	GAGTGATTGA	[181]
49_Strombomonas_triquetra	GATACTGTCC	TTTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTATAT	CCTCGCAAGG	CAGCATACTG	GGACCGATAG	AGCACAAAGTA	GAGTGATCGA	[196]
50_Eutreptiella_pomquetensis	GTATAGTCCC	CGTGAGTAAG	CTGCTTGGGA	CTGTAGTTTG	AACGGTGTAT	CCTCGTAAGG	CTGAATACCG	GGACCGATAG	AGTACAAGTA	GAGTGATCGA	[200]
51_Trachelomonas_planctonica	GATACTGTCC	CATGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	AAGCATACTG	GGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[196]
52_Trachelomonas_lefevrei	GATACTGTCC	CATGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	AAGCATACTG	GGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[198]
53_Trachelomonas_volvocina	GATACTGTCT	CATGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	AAGAACAACG	GGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[200]
54_Trachelomonas_conradi	GTATCTGTCC	CGTGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	CAGCATACTG	GGACCGATAG	CGCACAAAGTA	GAGTGATTGA	[182]
55_Trachelomonas_similis	GAGACTGTCT	CATGAGTAAG	TGGCTTGGGA	CTGCAGTTTG	AAGGGTGTAT	CCTCGCAAGG	AAGAACAACG	GGACCGATAG	TGCACAAGTA	GAGTGATTGA	[200]

	210	220	230	240	250	260	270	280	290	300	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----ACCCG	TCTTGAAACA	CGGACCAAGG	AGTCAAACAG	ACGCGCGGAG	[45]
02_Blastocrithidia_culicis	AAGtTgAaAa	GcAcTtTgGa	AaGaGaGTGa	CaAgAaCcTg	AaGtCgTgCa	TcGg-aCcCg	TcTtGaAaCa	CgGaCcAaGg	AgTcAaAcAg	AcGcGcGtGg	[131]
03_Herpetomonas_muscarum	AAGTTGAAAA	GCACTTTGGa	AAGAGAGT-G	CAAGAACCTG	AAGTCGTGTC	AAAA-ACCCG	TCT-----T	CGG-ACAAGG	AGTCAAACAG	ACGCGCGTGG	[124]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
05_Leishmania_infantum	AAGTTGAAAA	GCACTTTGGa	AAGAGAGTGA	AAAGAACCTG	AAGTCGTGGG	GCAA-ACCCG	TCTTGAAACA	CGGACCAAGG	AGTCAAACAG	ACGCGCGTGG	[293]
06_Leptomonas_bifurcata	AAGTTGAAAA	GCACTTTGGa	AAGAGAGTGA	AAAGAACCTG	AAGTCGTGTC	AAAA-ACCCG	TCTTGAAACA	CGGACCAAGG	AGTCAAACAG	ACGCGCGTGG	[131]
07_Leishmania_major	AAGTTGAAAA	GCACTTTGGa	AAGAGAGTGA	AAAGAACCTG	AAGTCGTGCA	AGAA-ACCCG	TCTTGAAACA	CGGACCAAGG	AGTCAAACAG	ACGCGCGTGG	[293]
08_Leishmania_tarentolae	AAGTTGAAAA	GCACTTTGGa	AAGAGAGTGA	AAAGAACCTG	AAGTCGTGCA	AGAA-A-CCG	TCTTGAAACA	CGGACCAAGG	AGTCAAACAG	ACGCGCGTGG	[292]
09_Trypanosoma_cruzi	AAGTTGAAAA	GCACTTTGGa	AAGAGAGTGA	CAAGAACCTG	AAGTCGTGCA	AGAA-ACCCG	TCTTGAAACA	CGGACCAAGG	AGTCAAACAG	ACGCGCGTGG	[293]
10_Trypanosoma_brucei	AAGTTGAAAA	GTACTTTGGA	AAGAGAGTGA	CAAGAACCTG	AAATCGTGTC	ATGG-ACCCG	TCTTGAAACA	CGGACCAAGG	AGTCAAACAG	ACGCGCCCGG	[293]
11_Bodo_saltans	AAGTTGAAAA	GCACTTTGGa	AAGAGAGTGA	CAAGAACCTG	AAGTCGTGTC	-CAA-ACCCG	TCTTGAAACA	CGGACCAAGG	AGTCAAACAG	ACGCGCGTGG	[228]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	AAGATGAAAA	GCACTTTGGa	AAGAGAAAT-C	AAAGTACGTG	AAACTGTGGC	TTCT-ACCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATCGT	ACCTGCAAGG	[292]
15_Colacium_elongatum	AGGATGCAAA	GGACTTTGGA	AAGAGGGT-T	AAAGTCCTGT	AAGCCGCGTC	CAAGTGCCCG	TCTYGAAACA	CGGACCAAGG	AGTCTACGGA	GCTTGCCAGG	[299]
16_Cryptoglena_pigra	AGGATGCAAA	GGACTTTGGA	AAGAGGGT-T	AAAGTCCTGT	AAGCTGTGCC	GGAAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAAAG	[298]
17_Cryptoglena_sukjai	AGGATGCAAA	GGACTTTGGA	AAGAGGGT-T	AAAGTCCTGT	AAGCTGTGCC	TGAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAAAG	[284]
18_Colacium_sp	AGGATGCAAA	GGACTTTGGA	AAGAGGGT-T	AAAGTCCTGT	AAGCCGCGTC	ACTGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[300]
19_Colacium_vesiculosum	AGGATGCAAA	GGACTTTGGA	AAGAGGGT-T	AAAGTCCTGT	AAGCCGCGTC	CAAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[300]
20_Euglena_clara	AAGGTCCATG	GGACCCGAAA	GGTGGCGAAC	TACCATCGCT	ACGACGAACA	CCTAGGCCCG	TCGGAGTCC	TGCCCCGTGT	GGTCTTTCA-	ACAGGCTGGT	[292]
21_Discoploastis_adunca	AGGATGCAAA	GGACTTTGGA	AAGAGGGT-T	AAAGTCCTGT	AAGCTGCAAG	TGGCTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCCTGG	[297]
22_Euglena_deses	AGGATGCAAA	GGACTCCGCG	AGGAGGGTTA	A-AGTCCTGT	AAGCCTCGGG	GGGGTGCCCG	TCTCGAAACA	CGGACCAAGG	AGTCTATGGA	GCCTGCCAGG	[297]
23_Euglena_geniculata	AGGATGCAAA	GAACCTCAGC	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCGCC	TCTGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCGGG	[299]
24_Euglena_granulata	AGGATGCAAA	GAACCTCCGCG	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCGAG	CAAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGA	[300]
25_Euglena_hiemalis	AGGATGCAAA	GAACCTCCGCG	AAGAGGGTTA	A-AGCCCTGT	AAGCCGCAATC	TTCC-GCCCG	TCTCGAAACA	CGGACCAAGG	AGTCTAAGGA	GCTTGCCAGG	[296]
26_Eutreptia_pertyi	ATGGTGCAAA	GTACTCTGCG	GAGAGGGTTT	A-AGTTCCCG	AAGTCGTACT	GTCCTGCCCG	TCTTGAAACA	CGGATCAAGG	AGTCTATGGA	ATGCGCCAGG	[298]
27_Euglena_stellata	ATGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AACGTCCCTG	AAGCCTCGCC	CTGCTACCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[285]
28_Euglena_longa	AGGATGCAAA	GAACCTCCGCG	AAGAGGGTTA	A-AGCCCTGT	AAGCCGCAAG	AAGC-GCCCG	TCTCGAAACA	CGGACCAAGG	AGTCTAAGGA	GCTTGCCAGG	[283]
29_Euglena_viridis	AGGATGCAAA	GAACCTCCGCG	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCATT	AAGGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[298]
30_Euglena_exilis	AGGATGCAAA	GGACTCTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCGCC	ACCG-GCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCCTGCCAGG	[299]
31_Euglena_gracilis	AGGATGCAAA	GAACCTCCGCG	AAGAGGGTTA	A-AGTCCTGT	AAGCCGCAATC	AAGCTGCCCG	TCTCGAAACA	CGGACCAAGG	AGTCTAAGGA	GCTTGCCAGG	[299]
32_Lepocinclis_steinii	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCACC	AAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[281]
33_Lepocinclis_acus_var_major	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCGCGCC	AGAG-GCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCCTGG	[282]
34_Lepocinclis_spirogyroides	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTTGCC	GGCGTGCCCG	TCTYGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCGGG	[283]
35_Lepocinclis_triapteris	AGGGTGCAAA	GGACTTTGCA	AAGAGGGTTT	AAAGTCCTGT	AAGCCTCGCC	GACGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCCTCG	[285]
36_Lepocinclis_ovum	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTTGTC	CAGG-GCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATAGA	GCTTGCCAGG	[284]
37_Monomorphina_aenigmatica	AGGATGCAAA	GGACTTTGCA	AAGAGGGT-T	AAAGTCCTGT	AAGCCGCACT	CTGTTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCCCGCCAAG	[299]
38_Colacium_mucronatum	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	A-AGTCCTGT	AAGCCGCGCC	TCAAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTACGGA	GCTTGCCAGG	[299]
39_Monomorphina_sp	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	A-AGTCCTGT	AAGCTGTGCC	TTAGTGCCCG	TCTTGAAACA	CGGATCAAGG	AGTCTATGGA	GCTTGCCAAG	[298]
40_Peranema_trichophorum	AATGTGAAAA	GCACCTGTGA	AAGAGAGT-T	AAAGAACCTG	AAGCCAAGCG	GACT-TCCCG	TCTTGTAACA	CGGATCAAGG	AGTCTACTGT	CAGTGCGAGT	[290]
41_Phacus_orbicularis	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCGCC	TGAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCGGG	[299]
42_Phacus_similis	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCACC	AGTGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCCTGG	[279]
43_Monomorphina_pyrum	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCTGTGCC	AGCATGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAAAG	[285]
44_Strombomonas_napiformis	AGGATGCAAA	GGACTCTGCA	AAGAGGGTTA	AAAGCCCTGT	AAGCCTCAGC	CCGCTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[296]
45_Strombomonas_borsteniensis	AGGATGCAAA	GGACTCTGCA	AAGAGGGTTA	A-AGCCCTGT	AAGCCTCAAG	TGAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[296]
46_Strombomonas_acuminata	AGGATGCAAA	GGACTCTGCA	AAGAGGGTTA	AAAGCCCTGT	AAGCCTCACC	AGAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[282]
47_Strombomonas_ovalis	AGGATGCAAA	GGACTCTGCA	AAGAGGGTTA	AAAGCCCTGT	AAGCCTCACC	AGCGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[297]
48_Trachelomonas_echinata	AGGATGCAAA	GCACCTTGGa	AAGGGGGTTC	AACGTCCCTG	AAGCCTCATC	CAAG-GCCCG	TCTTGAAACA	CGGATCAAGG	AGTCTATGGA	GCAGGCCAGG	[280]
49_Strombomonas_triquetra	AGGATGCAAA	GGACTCTGCA	AAGAGGGTTA	A-AGCCCTGT	AAGCCTCACC	AGCGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[295]
50_Eutreptiella_pomquetensis	AGGGTGCAAA	GAACCTGCA	AAGAGGGTTA	AAAGTACCTG	AAGCTGCAAC	CAAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGG	GCAGGCCGGG	[300]
51_Trachelomonas_planctonica	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCGTC	AAAC-GCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[295]
52_Trachelomonas_lefevrei	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTTGGT	AGAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[298]
53_Trachelomonas_volvocina	AGGATGCAAA	GTACTCTGCG	AAGAGAGTTA	AAAGTGCTGT	AAGCCTCACC	AGAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCCTGG	[300]
54_Trachelomonas_conradi	AGGATGCAAA	GGACTTTGCA	AAGAGGGTTA	AAAGTCCTGT	AAGCCTCGCC	AGCGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCAGG	[282]
55_Trachelomonas_similis	AGGATGCAAA	GTACTCTGCG	GAGAGAGTTA	AAAGTGCTGT	AAGCCTCACC	AGAGTGCCCG	TCTTGAAACA	CGGACCAAGG	AGTCTATGGA	GCTTGCCCTGG	[300]

	310	320	330	340	350	360	370	380	390	400	
01_Rhynchomonas_nasuta	GTT--GAGTT	AGTGTGCCCTG	TTTGGACCCG	AAAGGTGGTG	AACATATGCNT	GAACAAGATG	AAGCCCAGGG	AAACCCAGGT	GGAGGCTCTGT	AGCGGTGCTG	[143]
02_Blastocrithidia_culicis	GtT--gCgTt	AgTgTgCtTg	TtTgGaCcCg	AaAgGtGgTg	AaCtATGcCt	GaACaAgGtG	AaGcCcGgGg	AaAcCcAgGt	GgAgGcCtGt	AgCgAtGcTg	[229]
03_Herpetomonas_muscarum	GTT--GC6TT	AGTGTGCCCTG	TTTGGACC-C	AAAGGTGGTG	AACATATGCCT	GAACAAGGTC	AAGCCCAGGG	AAACCCAGGT	CGCAGCCTGT	AGCGATGCTG	[221]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
05_Leishmania_infantum	GTT--GC6TT	AGTGTGCCCTG	TTTGGACCCG	AAAGGTGGTG	AACATATGCCT	GAACAAGGTC	AAGCCCAGGG	AAACCCAG-G	GGAGGCTCTGT	AGCGATGCTG	[390]
06_Leptomonas_bifurcata	GTT--GC6TT	AGTGTGCCCTG	TTTGGACCCG	AAAGGTGGTG	AACATATGCCT	GAACAAGGTC	AAGCCCAGGG	AAACCCAGGT	GGAGGCTCTGT	AGCGATGCTG	[229]
07_Leishmania_major	GTT--GC6TT	AGTGTGCCCTG	TTTGGACCCG	AAAGGTGGTG	AACATATGCCT	GAACAAGGTC	AAGCCCAGGG	AAACCCAGGT	GGAGGCTCTGT	AGCGATGCTG	[391]
08_Leishmania_tarentolae	GTT--GC6TT	AGTGTGCCCTG	TTTGGACCCG	AAAGGTGGTG	AACATATGCCT	GAACAAGGTC	AAGCCCAGGG	AAACCCAGGT	GGAGGCTCTGT	AGCGATGCTG	[389]
09_Trypanosoma_cruzi	GTT--GAGTT	AGTGTGCCCTG	TTTGGACCCG	AAAGGTGGTG	AACATATGCCT	GAACAAGGTC	AAGCCCAGGG	AAACTCGGGT	GGAGGCTCTGT	AGCGATGCTG	[391]
10_Trypanosoma_brucei	-CC--CCCCG	AGTGTGCCCTG	TTTGGACCCG	AAAGGTGGTG	AACATATGCCT	GAACAAGGTC	AAGCCCAGGG	AAACCCCGGT	GGAGGCTCTGT	AGCGATGCTG	[390]
11_Bodo_saltans	GTT--GAGTT	AGTGTGCCCTG	TTTGGACCCG	AAAGGTGGTG	AACATATGCCT	GAACAAGGTC	AAGCCCAGGG	AAACCCAGGT	GGAGGCTCTGT	AGCGATGCTG	[326]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	TGGT-GAGG-	AGCAGGTGCG	ATGGGACCCG	AAAGATGGTG	ATCTTTTCT	GAATAGGGCG	AAGGCGGGG	AAACCTGCT	GGAGGCTCTGT	AGCGCTACTG	[390]
15_Colacium_elongatum	GTTTTGAGTG	TGCAAGGTCC	GTGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTATGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[399]
16_Cryptoglena_pigra	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGATGGTG	AACATATACCA	TTCTATGGTG	AAGCCGTGGG	AAACTGCGGT	GGAAGCCAGT	AGCTGTTCTG	[398]
17_Cryptoglena_sukjai	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGATGGTG	AACATATACCA	TTCTATGGTG	AAGCCGTGGG	AAACTGCGGT	GGAAGCCAGT	CGCTGTTCTG	[384]
18_Colacium_sp	GTTTTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTATGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[400]
19_Colacium_vesiculosum	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTATGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[400]
20_Euglena_clara	TT-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[294]
21_Discoplastis_adunca	GTTTTGAGT-	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCT	GACTACAACG	AAGCCCTGGG	AAACTGGGGT	GGAGGTTGGT	CGCTGTTCTG	[396]
22_Euglena_deses	GTTTTGAGTG	TGCGCGGTCC	AAGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTACGACG	AAGCCATGGG	AAACTGTGGT	GGAAGTCGGT	CGCTGTTCTG	[397]
23_Euglena_geniculata	GTTTTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATGCCA	TTCCACACCG	AAGCCATGGG	AAACTGTGGT	GGAAGTGGT	AGCTGTTCTG	[399]
24_Euglena_granulata	GTTTCGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCA	TTCTACGACG	AAGCCATGGG	AAACTGTGGT	GGAAGTCGGT	CGCTGTTCTG	[400]
25_Euglena_hiemalis	GTTTTGAGTG	TGCAAGGTCC	TTGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTACGGTG	AAGCCATGGG	AAACTGTGGT	GGAGGCCGGT	AGCTGTTCTG	[396]
26_Eutryptia_pertyi	GTTCTGAGTG	AGCGTCCGCG	ATAGGACCCG	AAAGATGGTG	AACATATACCT	GTTTACGATG	AAGCTGTGGG	AAACTGCTGT	GGAGGTCGGA	CAGTGTCCTG	[398]
27_Euglena_stellata	GTTTTGAGTG	TGCAATTTC	ATGGGACCCG	AAAGGTGGTG	AACATATACCT	TTCTACAACG	AAGCCATGGG	AAACTGTGGT	GGAAGTTGGT	CGCTGTTCTG	[385]
28_Euglena_longa	GTTTTGAGTG	TGCAAGGTCC	TTGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTACGATG	AAGCCATGGG	AAACTGTGGT	GGAAGTCGGT	CGCTGTTCTG	[383]
29_Euglena_viridis	GTTTTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTACGACG	AAGCCATGGG	AAACTGTGGT	GGAAGTCGGT	CGCTGTTCTG	[398]
30_Euglena_exilis	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCA	TTCCACGCGG	AAGCCCTGGG	AAACCCGGGT	GGAGGCCGGT	CGCTGTTCTG	[399]
31_Euglena_gracilis	GTTTCGAGTG	TGCAAGGTCC	TTGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTACGGTG	AAGCCATGGG	AAACTGTGGT	GGAAGCCGGT	CGCTGTTCTG	[399]
32_Lepocinclis_steinii	GTTTTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TACCATGATG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[381]
33_Lepocinclis_acus_var_major	GTTTTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCA	TACCATGACG	AAGCCATGGG	AAACTGTGGT	GGAGGTCAGT	CGCTGTTCTG	[382]
34_Lepocinclis_spirogyroides	GTTTTGAGTG	TGCGAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TACCATGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCGGC	CGCTGTTCTG	[383]
35_Lepocinclis_triptyeris	GTTTTGAGT-	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCCACGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCGGT	CGCTGTTCTG	[384]
36_Lepocinclis_ovum	GTTTTGAGTG	TGCGAGGTCT	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TACCATGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCGGT	CGCTGTTCTG	[384]
37_Monomorphina_aenigmatica	GTTCTGAGTG	TGCGTGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCA	TTCTATGGTG	AAGCCTTGGG	AAACCCAGGT	GGAGGCCAGT	AGCTGTTCTG	[399]
38_Colacium_mucronatum	GTTTTGAGTG	TGCAAGGTCC	GTGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTATGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[399]
39_Monomorphina_sp	GTTCTGAGTG	TGCAAGGTCC	CTGGGACCCG	AAAGGTGGTG	AACATATACCA	TTCTATGGTG	AAGCCTTGGG	AAACTGAGGT	GGAGGCCAGC	AGCTGTTCTG	[398]
40_Peranema_trichophorum	TTT--GAGT-	AGCATTGGTA	GTTGGACCCG	AAAGGTGGTG	AGCTATACCC	TTGGACAGTG	AAGTCATGGG	AAACTGTGAT	GGAAGTCGGT	TACTGTGCTA	[387]
41_Phacus_orbicularis	GTTTTGAGTG	GGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCT	AGCCACGACG	AAGCCATGGG	AAACTGTGGT	GGAGGTCGGT	CGCTGTTCTG	[399]
42_Phacus_similis	GTTTTGAGT-	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCCATGACG	AAGCCATGGG	AAACTGTGGT	GGAGGTCAGT	CGCTGTCCTG	[378]
43_Monomorphina_pyrum	GTTCTGAGTG	TGCAAGGTCC	CTGGGACCCG	AAAGGTGGTG	AACATATACCG	TGCTATGATG	AAGCCTTGGG	AAACTGAGGT	GGAGGCCAGC	AGCTGTTCTG	[385]
44_Strombomonas_napiformis	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TGCTATGATG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[396]
45_Strombomonas_borysteniensis	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TGCTATGATG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[396]
46_Strombomonas_acuminata	GTTCTGAGT-	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TGCTATGATG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[381]
47_Strombomonas_ovalis	GTTCTGAGT-	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TGCTATGATG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[396]
48_Trachelomonas_echinata	GTTCTGAGTG	TGCAATGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TGCCATGGTG	AAGCTGTGGG	AAACTGCGGT	GGAGGCCAGA	AGCTGTTCTG	[380]
49_Strombomonas_triquetra	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TGCTATGATG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[395]
50_Eutreptiella_pomquetensis	GTTCTGAGTG	AGCGTGTGCC	GTGGGACCCG	AAAGGTGGTG	AACATATACCT	GGTCACGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCGGT	AGCTGTCCTG	[400]
51_Trachelomonas_planctonica	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTATGTCG	AAGCCGTGGG	AAACTGCGGT	GGAGGACAGT	CGCTGTTCTG	[395]
52_Trachelomonas_lefevrei	GTTCTGAGTG	TGCAATGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTATGTCG	AAGCCGTGGG	AAACTGCGGT	GGAGGACAGT	CGCTGTTCTG	[398]
53_Trachelomonas_volvocina	GTT--CGGAG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTATGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[398]
54_Trachelomonas_conradi	GTTCTGAGTG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTACGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCGGT	CGCTGTTCTG	[382]
55_Trachelomonas_similis	GTT--CGGAG	TGCAAGGTCC	ATGGGACCCG	AAAGGTGGTG	AACATATACCG	TTCTATGACG	AAGCCGTGGG	AAACTGCGGT	GGAGGTCAGT	CGCTGTTCTG	[398]

	410	420	430	440	450	460	470	480	490	500	
01_Rhynchomonas_nasuta	ACGTGCAAA	CGCT-CGCAT	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACATCGAAGT	TTCCCTCAGG	ATAGCTGGTG	[242]
02_Blastocrithidia_culicis	AcGtGcAAAt	CgCt-cGtAg	GaTtTgGgTa	TcGgAgCgAa	AgAcTcAtCg	AaCcAcCtAg	TaGcTgGtTc	AcAtCgAaGt	TtCcCtCaGg	AtAgCtGgTg	[328]
03_Herpetomonas_muscarum	ACGTGCAA-A	CGCT-CGTAT	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACATCGAAGT	T-TCCTCAGG	-AAGCTGGTG	[317]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
05_Leishmania_infantum	ACGTGCAAA	CGCT-CGTAT	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACATCGAAGT	TTCCCTCAGG	ATAGCTGGTG	[489]
06_Leptomonas_bifurcata	ACGTGCAAA	CGCT-CGTAT	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACATCGAAGT	TTCCCTCAGG	ATAGCTGGTG	[328]
07_Leishmania_major	ACGTGCAAA	CGCT-CGTAT	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACATCGAAGT	TTCCCTCAGG	ATAGCTGGTG	[490]
08_Leishmania_tarentolae	ACGTGCAAA	CGCT-CGTAT	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACATCGAAGT	TTCCCTCAGG	ATAGCTGGTG	[488]
09_Trypanosoma_cruzi	ACGTGCAAA	CGCT-CGCAT	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACATCGAAGT	TTCCCTCAGG	ATAGCTGGTG	[490]
10_Trypanosoma_brucei	ACGTGCAAA	CGCT-CGTAC	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACGTGGAAGT	TTCCCTCAGG	ATAGCTGGTG	[489]
11_Bodo_saltans	ACGTGCAAA	CGCT-CGCAC	GATTTGGGTA	TCGGAGCGAA	AGACTCATCG	AACCACCTAG	TAGCTGGTTC	ACATCGAAGT	TTCCCTCAGG	ATAGCTGGTG	[425]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	ACGTGCAAA	CGTT-CGTTA	AATTGCGGTA	AAGGGGCGAA	AGACTCATCG	AACCTGTCTAG	TAGCTGGTTC	CCTCCGAAGT	TTCCCTCAGG	ATAGCAGGAG	[489]
15_Colacium_elongatum	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGG-	[498]
16_Cryptoglena_pigra	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCATCTAG	TAGCTGGTTT	CCCTTGAAGT	TTCCCTCAGG	ATAGCTGGGT	[498]
17_Cryptoglena_sukjai	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCATCTAG	TAGCTGGTTT	CCCTTGAAGT	TTCCCTCAGG	ATAGCTGGGT	[484]
18_Colacium_sp	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGG-	[499]
19_Colacium_vesiculosum	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[500]
20_Euglena_clara	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[294]
21_Discoplastis_adunca	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[496]
22_Euglena_deses	ACGTGCAAA	CGTGTAGCCT	GAGCGCGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[497]
23_Euglena_geniculata	ACGTGCAAA	CGTGTAGCCC	GAGAGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[499]
24_Euglena_granulata	ACGTGCAAA	CGTGTAGCCT	GAGAGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTAAGG	ATAGCTGGGC	[500]
25_Euglena_hiemalis	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGA	[496]
26_Eutreptia_pertyi	ACGTGCAAA	CGAGTTGTTG	GAATGAGGCA	TAGGGGTGAA	AGGCCAATCG	AACCGTCTAG	TAGCTGGTTT	CATCCGAAGT	TTCCCTCAGG	ATAGCCGAGA	[498]
27_Euglena_stellata	ACGTGCAAA	CGTGTAGCCT	GAGCGCGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[485]
28_Euglena_longa	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGA	[483]
29_Euglena_viridis	ACGTGCAAA	CGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[498]
30_Euglena_exilis	ACGTGCAAA	TGTGTAGCCT	GAGCGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[499]
31_Euglena_gracilis	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGA	[499]
32_Lepocinclis_steinii	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[481]
33_Lepocinclis_acus_var_major	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCACTGAAGT	TTCCCTCAGG	ATAGCTGGGC	[482]
34_Lepocinclis_spirogyroides	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[483]
35_Lepocinclis_triapteris	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCATCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[484]
36_Lepocinclis_ovum	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGT	[484]
37_Monomorphina_aenigmatica	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCCCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[499]
38_Colacium_mucronatum	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[499]
39_Monomorphina_sp	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[498]
40_Peranema_trichophorum	ACGTGCAAA	TGCT-AGTCA	CACATGGGCA	TAGGGGCGAT	AGGCTCATCG	AACCACCTAG	TAGCTGGTTT	CTTCTGAAGT	TTCCCTCAGG	ATAGCTGGAAG	[486]
41_Phacus_orbicularis	ACGTGCAAA	TGTGTAGCCT	GAGCTTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[499]
42_Phacus_similis	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGT	[478]
43_Monomorphina_pyrum	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[485]
44_Strombomonas_napiformis	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCAGGGT	[496]
45_Strombomonas_borysteniensis	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCCAGGT	[496]
46_Strombomonas_acuminata	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCCGGGT	[481]
47_Strombomonas_ovalis	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCCAGGT	[496]
48_Trachelomonas_echinata	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTC	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGG-	[479]
49_Strombomonas_triquetra	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCCGGGT	[495]
50_Eutreptiella_pomquetensis	ACGTGCAAA	TGTGTAGTCT	GAATTGGGCA	TAGGGGTGAA	AGGCCAATCG	AACCGTCTAG	TAGCTGGTTT	CCACCGAAGT	TTCCCTCAGG	ATAGCCGGGG	[500]
51_Trachelomonas_planctonica	ACGTGCAAA	TGTGTAGCCT	GAGGCGGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCTTCGAAGT	TTCCCTCAGG	ATAGCTGGGT	[495]
52_Trachelomonas_lefevrei	ACGTGCAAA	TGTGTAGCCT	GAGGCGGGCA	TAGGGGTGAA	AGGCCAATCG	AACCACCTAG	TAGCTGGTTT	CCTT-CGAAGT	TTCCCTCAGG	ATAGCTGGGT	[497]
53_Trachelomonas_volvocina	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[498]
54_Trachelomonas_conradi	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGT	[482]
55_Trachelomonas_similis	ACGTGCAAA	TGTGTAGCCT	GAGTGTGGCA	TAGGGGTGAA	AGGCCAATCG	AGCCACCTAG	TAGCTGGTTT	CCCTCGAAGT	TTCCCTCAGG	ATAGCTGGGC	[498]

	510	520	530	540	550	560	570	580	590	600	
01_Rhynchomonas_nasuta	CTGTAAAGCA	AATGATTAGA	GGCAGATATC	GACCTATTCT	CAAACCTTTAA	ATGTGCAAAAC	TTGT-G----	-----	CGGAGCAGGC	GATGCGGAAT	[327]
02_Blastocrithidia_culicis	CtGtAaAgCa	AaTgAtTaGa	GgCaAaTaTc	GaCcTaTtCt	CaAaCtTtAa	AtGtGcAaAt	CcAgG-----	-----	CgGaGcAgGc	GaTgCgGaAt	[413]
03_Herpetomonas_muscarum	CTGTAAAGCA	AATGATTAGA	GGCAAAATATC	GACCT-ATCT	CAAACCTT-TA	ACAAACAAAC	AATCGG----	-----	CGGAGGAGGC	GATGCGGAAT	[401]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
05_Leishmania_infantum	CTGTAAAGCA	AATGATTAGA	GGCAAAATATC	GACCTATTCT	CAAACCTTTAA	ATGTGCAAAAC	AACTGC----	-----	CGGAGCAGGC	GATGCGGAAT	[575]
06_Leptomonas_bifurcata	CTGTAAAGCA	AATGATTAGA	GGCAAAATATC	GACCTATTCT	CAAACCTTTAA	ATGTGCAAAAC	AACTGC----	-----	CGGAGCAGGC	GATGCGGGAAT	[414]
07_Leishmania_major	CTGTAAAGCA	AATGATTAGA	GGCAAAATATC	GACCTATTCT	CAAACCTTTAA	ATGTGCAAAAC	AACTGC----	-----	CGGAGCAGGC	GATGCGGAAT	[576]
08_Leishmania_tarentolae	CTGTAAAGCA	AATGATTAGA	GGCAAAATATC	GACCTATTCT	CAAACCTTTAA	ATGTGCAAAAC	AACTGC----	-----	CGGAGCAGGC	GATGCGGAAT	[574]
09_Trypanosoma_cruzi	CTGTAAAGCA	AATGATTAGA	GGCAAAATATC	GACCTATTCT	CAAACCTTTAA	ATGTGCAAAAT	TTAGC-----	-----	CGGAGCAGGC	GATGCGGAAT	[575]
10_Trypanosoma_brucei	CTGTAAAGCA	AATGATTAGA	GGCAAAATATC	GACCTATTCT	CAAACCTTTAA	ATGTGCAAAAG	TTATTT----	-----	CGGAGCAGGC	GATGCGGAAT	[575]
11_Bodo_saltans	CTGTAAAGCA	AATGATTAGA	GGCAAAATATC	GACCTATTCT	CAAACCTTTAA	ATGTGCGAGAT	AATCGC----	-----	CGGAGCAGGC	GATGCGGAAT	[511]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	TGGTAAAGCG	AATGATTAGA	GGCTTTCTCT	GACCTATTCT	CAAACCTTTAA	ATGGGCAAGT	TAATT-----	-----	CAGGACTGGC	GATGCGGGAT	[574]
15_Colacium_elongatum	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	GACCTATTCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAACTG	TGTTGTTCTC	CAGTGCTGGC	GATGCGGGAT	[598]
16_Cryptoglena_pigra	TCGTAAAGCA	ACGCATGAGA	GGCATTCCTC	AACCTATTCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAACTG	GTTGGTCTCT	CAGTGCTGGT	GATGCGGAAT	[598]
17_Cryptoglena_squajai	TCGTAAAGCA	ACGCATGAGA	GGCACTGCC	AACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GCTGGTCTCT	CAGTGCTGGT	GATGCGGAAT	[584]
18_Colacium_sp	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	GACCTATTCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAACTG	TGCTGTTCTC	CAGTGCTGGC	GATGCGGGAT	[599]
19_Colacium_vesiculosum	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	GACCTATTCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAACTG	GCTGGTCTCC	CAGTGCTGGC	GATGCGGGAT	[600]
20_Euglena_clara	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[323]
21_Discoplastis_adunca	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTT	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	ACTTCACTGA	TGCTTGTCTC	CAGTGCTGGC	GACGCGGGAT	[596]
22_Euglena_deses	TCGTAAAGCA	ACGCATGAGA	GGCCCTTCTC	AACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GCCAGTTCTT	CAGTGCTGGT	GATGCGGGAT	[597]
23_Euglena_geniculata	TCGTAAAGCA	ACGCATGAGA	GGCGCTTCTC	CACCTATTCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAACTG	TGCCCTGTTT	CAGTGCTGGC	GACGCGGGAT	[598]
24_Euglena_granulata	TCGTAAAGCA	ACGCATGAGA	GGCACCCTCT	GACCTATCCT	CGAACTGTGA	ATGGGCGTGT	CCTTCGC---	-----	CAGTGCTGGC	GATGCGGGAT	[587]
25_Euglena_hiemalis	TCGTAAAGCA	ACGCATGAGA	GGCCCTCCTC	AACCTATCCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GACTGTTCTC	CAGTGCTGGT	GATGCGGGAT	[596]
26_Eutreptia_pertyi	TCGTAAAGCA	ACGCATGAGA	GGT-TTCTC	GACCTATTCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAAA-T-	-----	CAGTGCTGGC	GATGCGGGAT	[585]
27_Euglena_stellata	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	AACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TC-ACAACG-	CCTGGTCTCC	CAGTGCTGGC	GATGCGGGAT	[583]
28_Euglena_longa	TCGTAAAGCA	ACGCATGAGA	GGCCCTCCTC	AACCTATCCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GACCGTTCTC	CAGTGCTGGT	GATGCGGGAT	[583]
29_Euglena_viridis	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	AACCTATTCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAACTG	GCTGTCTCTC	CAGTGCTGGC	GATGCGGGAT	[598]
30_Euglena_exilis	CCGTAAAGCA	AC-GATGAGC	GGTTCCTCTC	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCACCAACGC	CCGGGCCCGG	CAGTGCTGGT	GATGCGGGAT	[598]
31_Euglena_gracilis	TCGTAAAGCA	ACGCATGAGA	GGCCCTCCTC	AACCTATCCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GACAGTTCTC	CAGTGCTGGT	GATGCGGGAT	[599]
32_Lepocinclis_steinii	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	AACCTATCCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAACTG	TGCTTCTCTC	CAGTGCTGGT	GATGCGGGAT	[581]
33_Lepocinclis_acus_var_major	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTT	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	GCTTCAACTG	TGCTTGGTAC	CAGTGCTGGT	GATGCGGGAT	[582]
34_Lepocinclis_spirogyroides	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	CGCCGTCTCT	CAGTGCTGGT	GATGCGGGAT	[583]
35_Lepocinclis_triapteris	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	AACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCACCCCTAT	GTCCGGGGAC	CAGTGCTGGT	GATGCGGGAT	[584]
36_Lepocinclis_ovum	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTA	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCATCGATGC	CTGAGTCGTT	CAGTGCTGGT	GATGCGGGAT	[584]
37_Monomorphina_aenigmatica	TTGTAAAGCA	ACGCATGAGA	GGCACCCTCT	AACCTATTCT	CGAACTGTGA	ATGGGCAAGT	TCTTCAACTG	TCTGGAGTCC	CAGTGCTGGT	GATGCGGGAT	[598]
38_Colacium_mucronatum	TCGTAAAGCA	ACGCATGAGA	GGCAGTCCTC	AACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GCTGGTCTCC	CAGTGCTGGC	GATGCGGGAT	[599]
39_Monomorphina_sp	TTGTAAAGCA	ACGCATGAGA	GGCTCTCCTC	AACCTATTCT	CGAACTGTGA	ATGGGCGTGT	CCTTCAACTG	TGCTGGTCTC	CAGTGCTGGT	GATGCGGGAT	[598]
40_Peranema_trichophorum	CCGTAAAG-C	AATGTATGGG	AGGATGACTC	GGCCTTTCTT	CAAACCTGTGA	ATCA--GACT	TATTTCCATC	TTGAT----	CAGTACTGGC	GATGCGGAAT	[578]
41_Phacus_orbicularis	TCGTAAAGCA	ACGCATGAGA	GGCATCCCTC	GACCTATCCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	CAGGCCGGTC	CAGTGCTGGT	GATGCGGGAT	[599]
42_Phacus_similis	TCGTAAAGCA	ACGCATGAGA	GGCATCCCTC	AACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TC-A-CCTGT	GCTGGTCT-C	CAGTGCTGGC	GATGCGGGAT	[575]
43_Monomorphina_pyrum	TTGTAAAGCA	ACGCATGAGA	GGCTCTCCTC	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GCTGGCCCTC	CAGTGCTGGT	GATGCGGGAT	[585]
44_Strombomonas_napiformis	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	GACCTATCCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	TGCTGTTCTC	CAGTGCTGGT	GATGCGGGAT	[596]
45_Strombomonas_borsteniensis	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	AACCTATCCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	TGCTGTTCTC	CAGTGCTGGT	GATGCGGGAT	[596]
46_Strombomonas_acuminata	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	GACCTATCCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	TGCTGTTCTC	CAGTGCTGGT	GATGCGGGAT	[581]
47_Strombomonas_ovalis	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	AACCTATCCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	TGCTGTTCTC	CAGTGCTGGT	GATGCGGGAT	[596]
48_Trachelomonas_echinata	TTGTAAAGCA	ACGCATGAGA	GGCACTGCTC	AACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TC-ACC----	-----	CAGTGCTGGT	GATGCGGGAT	[564]
49_Strombomonas_triquetra	TCGTAAAGCA	ACGCATGAGA	GGCACTCCTC	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	TGCTGTTCTC	CAGTGCTGGC	GACGCGGGAT	[595]
50_Eutreptiella_pomquetensis	CCGTAAAGCA	ACGCATGAGA	GGCTCTCCTC	GACCTATTCT	CAAACCTGTGA	ATGGGCAAGT	TC-ACAC-CT	ACGGCCGACT	CAGTGCTGGC	GATGCGGGAT	[598]
51_Trachelomonas_planctonica	TCGTAAAGCA	ACGCATGAGA	GGCGCTCCTC	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	TGCTGTTCTC	CAGTGCTGGT	GATGCGGGAT	[595]
52_Trachelomonas_lefevrei	TCGTAAAGCA	ACGCATGAGA	GGCGCTCCTG	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	TGCTGTTCTC	CAGTGCTGGT	GATGCGGGAT	[597]
53_Trachelomonas_volvocina	TCGTAAAGCA	ACGCATGAGA	GGCACTCTCT	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GCTGGTCTCC	CAGTGCTGGT	GATGCGGGAT	[598]
54_Trachelomonas_conradi	TCGTAAAGCA	ACGCATGAGA	GGCTCTCCTG	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GCTGGTCTCC	CAGTGCTGGT	GATGCGGGAT	[581]
55_Trachelomonas_similis	TCGTAAAGCA	ACGCATGAGA	GGCACTCTCT	GACCTATTCT	CGAACTGTGA	ATGGGCGTGT	TCTTCAACTG	GCCGGTCTCC	CAGTGCTGGT	GATGCGGGAT	[598]

	610	620	630	640	650	660	670	680	690	700	
01_Rhynchomonas_nasuta	GAACCGCTTA	GGGATATGTG	CGGTCAAAC	TTGGGCTCAT	T-CTCGATAC	GTGAAAGGGT	TTTGGTGGAT	TAAGACAGTT	TGACGGTGGT	CATGGAAGTC	[426]
02_Blastocrithidia_culicis	GaACcGcTtTa	GgGaTaTgTg	GcGtCaAaCt	TtGgGcTcAt	T-CtCgAtAc	GtGaaGgGt	GtTgGtGgAt	-aGgAcAgTt	GgAcGgTgGc	CaTgGaAgTc	[511]
03_Herpetomonas_muscarum	GAACCGCTTA	GGGATATGTG	CGGTCAAAC	TTGGGCTCAT	--TCTC-ATC	GTGAAAGGGT	TTTGGTGGAT	ACGGACAGTT	GGACGGTGGC	CATGGAAGTC	[498]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
05_Leishmania_infantum	GAACCGCTTA	GGGATATGTG	CGGTCAAAC	TTGGGCTCAT	T-CTCGATAC	GTGAAAGGGT	TTTGGTGGAT	ATGGACAGTT	GGACGGTGGC	CATGGAAGTC	[674]
06_Leptomonas_bifurcata	GAACCGCTTA	GGGATATGTG	CGGTCAAAC	TTGGGCTCAT	T-CTCGATAC	GTGAAAGGGT	TTTGGTGGAT	ATGGACAGTT	GGACGGTGGC	CATGGAAGTC	[513]
07_Leishmania_major	GAACCGCTTA	GGGATATGTG	CGGTCAAAC	TTGGGCTCAT	T-CTCGATAC	GTGAAAGGGT	TTTGGTGGAT	ATGGACAGTT	GGACGGTGGC	CATGGAAGTC	[675]
08_Leishmania_tarentolae	GAACCGCTTA	GGGATATGTG	CGGTCAAAC	TTGGGCTCAT	T-CTCGATAC	GTGAAAGGGT	TTTGGTGGAT	ATGGACAGTT	GGACGGTGGC	CATGGAAGTC	[673]
09_Trypanosoma_cruzi	GAACCGCTTA	-GGATATGTG	GCGTCAAAC	TTGGGCTCAT	T-CTCGATAC	GTGAAAGGGT	TTTGGTGGAT	AGGGACAGTT	GGACGGTGGC	CATGGAAGTC	[673]
10_Trypanosoma_brucei	GAACCGCTTA	GGGATATGTG	CGGTCAAAC	TTGGGCTCAT	T-CTCGATAC	GTGAAAGGGT	TTTGGTGGAT	AGGGACAGTT	GGACGGTGGC	CATGGAAGTC	[674]
11_Bodo_saltans	GAACCGCTTA	GGGATATGTG	GCGTCAAAC	TTTGGCTCAT	T-CTCGATAC	GTGAAAGGGT	TTTGG-TGAT	A-TGACAGTT	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	GCTCCGAAG	CGTATTTAAG	GAATCCGAGC	GGTGGCTCAT	--T-AGATCC	CATGAAGGAT	GTTCGTTGAT	CTAGACAGCA	GGACGGTGGC	CATGGAAGTT	[671]
15_Colacium_elongatum	GCTCCGAGAG	CGGACTTAAA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTCCGCTGCT	AGAAATAGCA	GGACGGTGGC	CATGAAGGTT	[698]
16_Cryptoglena_pigra	GCTCCGAGAG	TGGATTTACA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGACGGTGGC	CATGAAGGTT	[698]
17_Cryptoglena_sujai	GCTCCGAGAG	TGGATTTACA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGACGGTGGC	CATGAAGGTT	[684]
18_Colacium_sp	GCTCCGAGAG	CGGACTTAAA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGAAATAGCA	GGACGGTGGC	CATGAAGGTT	[699]
19_Colacium_vesiculosum	GCTCCGAGAG	CAGACTTAAA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGACGGTGGC	CATGAAGGTT	[700]
20_Euglena_clara	GCTCCGAGAG	TGGTCTTACA	GAGCCCGAGT	GCTGGCTCAT	ACTCAGAACC	CACAAAGGAT	GTTCGCTGCT	CTGGATAGCA	GGACGGTGGC	CATGAAGGTT	[423]
21_Discoplistis_adunca	GCTCCGAGAG	TGGCCTTACA	GAACCCAGAGT	GCTGGCTCAT	ACACAGATCC	CACAAAGGAT	GTTCGCTGCT	AGTGATAGCA	GGACAGTGGC	CATGAAGGTT	[696]
22_Euglena_deses	GCTCCGAGAG	TGGACTTACA	GATCCCGAGT	GCTGGCTCAT	ACTCAGATCC	CACAAAGGAT	GTTCGCTGCT	CTGGATAGCA	GGACGGTGGC	CATGAAGGTT	[697]
23_Euglena_geniculata	GCTCCGAGAG	TGGATTTACA	GATCCAGAGT	GCTGGCTCAT	ACTCAGATCC	CACAAAGGAT	GTTCGCTGCT	CTTGATAGCA	GGACGGTGGC	CATGAAGGTT	[698]
24_Euglena_granulata	GCTCCGAGAG	CGGATCTACA	GAACCCGAGT	GCTGGCTCAT	ACTCAGATCC	CACAAAGGAT	GTTCGCTGCT	CTGGATAGCA	GGACGGTGGC	CATGAAGGTT	[687]
25_Euglena_hiemalis	GCTCCGAGAG	CAGACTTACA	GATCCAGAGT	GTTGGCTCAT	-ATCAGATCC	CACAAAGGAT	GTTCGCTTACT	ATGGATAGCA	GGACGGTGGC	CATGAAGGTT	[695]
26_Eutreptia_pertyi	GATCCGAAG	TGTGCTTACA	GAGTCCGAGT	GCTGGTTCAA	GCCCAGATCC	CACAAAGGAT	GTTCGCTTACT	ACTGATAGCA	GGACGGTGGC	CATGAAGGTT	[685]
27_Euglena_stellata	GCTCCGAGAG	CCGATTTACA	GATCCCAAGT	GCTGGCTCAT	ACTCAGATCC	CACAAAGGAT	GTTCGCTTACT	CTGGATAGCA	GGACGGTGGC	CATGAAGGTT	[683]
28_Euglena_longa	GCTCCGAGAG	CAGACTTACA	GATCCAGAGT	GTTGGCTCAT	-ATCAGATCC	CACAAAGGAT	GTTCGCTTACT	ATGGATAGCA	GGTCGGTGGC	CATGAAGGTT	[682]
29_Euglena_viridis	GCTCCNAGAG	TGGATTTACA	GATCCCGAGT	GCTGGCTCAT	ACTCAGATCC	CACAAAGGAT	GTTCGCTTACT	CTGGATAGCA	GGACGGTGGC	CATGAAGGTT	[698]
30_Euglena_exilis	GCTCCGAGAG	CGGCTTACA	GGGCCCGAGT	GCTGGCTCAT	ACACAGATCC	CACAAAGGAT	GTTCGCTGCT	AGAGATAGCA	GGGCGGTGGC	CATGAAGGTT	[698]
31_Euglena_gracilis	GCTCCGAGAG	CAGACTTACA	GATCCCGAGT	GCTGGCTCAT	-ATCAGATCC	CACAAAGGAT	GTTCGCTTACT	ACGGATAGCA	GGTCGGTGGC	CATGAAGGTT	[698]
32_Lepocinclis_steinii	GCTCCGAGAG	TGGACTTACA	GATCCCGAGT	GCTGGCTCAT	ACTAGATCC	CACAAAGGAT	GTTCGCTGCT	ATGGATAGCA	GGACGGTGGC	CATGAAGGTT	[681]
33_Lepocinclis_acus_var_major	GCTCCGAGAG	TGGACTTACA	GATCCCGAGT	GCTGGCTCAT	-ATCAGATCC	CACAAAGGAT	GTTCGCTGCT	ATGGATAGCA	GGTCAGTGGC	CATGAAGGTT	[681]
34_Lepocinclis_spirogyroides	GCTCCGAGAG	TGGACTTACA	GATCCAGAGT	GCTGGCTCAT	ACTCAGATCC	CACAAAGGAT	GTTCGCTGCT	ATGGATAGCA	GGACGGTGGC	CATGAAGGTT	[683]
35_Lepocinclis_triapteris	GCTCCGAGAG	TGGACTTACA	GATCCCGAGT	GCTGGCTCAT	ACACAGATCC	CACAAAGGAT	GTTCGCTGCT	CTGGATAGCA	GGTCGGTGGC	CATGAAGGTT	[684]
36_Lepocinclis_ovum	GCTCCGAGAG	TGGATTTACA	GATCCCGAGT	GCTGGCTCAT	ACACAGATCC	CACAAAGGAT	GTTCGCTGAT	TTGGATAGCA	GGTCGGTGGC	CATGAAGGTT	[684]
37_Monomorphina_aenigmatica	GCTCCGAGAG	TGGACTTACA	GGTCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGACGGTGGC	CATGAAGGTT	[698]
38_Colacium_mucronatum	GCTCCGAGAG	TGGACTTAAA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGACGGTGGC	CATGAAGGTT	[699]
39_Monomorphina_sp	GCTCCGAGAG	TGGACTTACA	GGTCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGAGATAGCA	GGACGGTGGC	CATGAAGGTT	[698]
40_Peranema_trichophorum	GCTCCGAGAG	TGTATGTAAA	GTGTCCAAAT	CTGAGCTCAT	--T-AGACAC	CC-TAAGGAT	GTTCGCTGCT	GAAGATAGCA	GGGCGGTGGC	CATGAAGGTT	[674]
41_Phacus_orbicularis	GCTCCGAGAG	CGGACTTACA	GTTCAGAGT	GCTGGCTCAT	-ATCAGATCC	CACAAAGGAT	GTTCGCTGCT	ATGGATAGCA	GGTCGGTGGC	CATGAAGGTT	[698]
42_Phacus_similis	GCTCCGAGAG	TGGACTTACA	GAACCTTAGT	GCTGGCTCAT	ACACAGATCC	CACAAAGGAT	GTTCGCTGCT	ATGGATAGCA	GGACGGTGGC	CATGAAGGTT	[675]
43_Monomorphina_pyrum	GCTCCGAGAG	TGGACTTACA	GGTCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGAGATAGCA	GGACGGTGGC	CATGAAGGTT	[685]
44_Strombomonas_napiformis	GCTCCGAGAG	TGGACTTGAA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGTGATAGCA	GGACGGTGGC	CATGAAGGTT	[696]
45_Strombomonas_borysteniensis	GCTCCGAGAG	TGGACTTGAA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGTGATAGCA	GGACGGTGGC	CATGAAGGTT	[696]
46_Strombomonas_acuminata	GCTCCGAGAG	TGGACTTGAA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGTGATAGCA	GGACGGTGGC	CATGAAGGTT	[681]
47_Strombomonas_ovalis	GCTCCGAGAG	TGGACTTGAA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGTGATAGCA	GGACGGTGGC	CATGAAGGTT	[696]
48_Trachelomonas_echinata	GCTCCGAGAG	TAGACTTACA	GGACAGAGT	GCTGGCTCAT	ACACAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGACGGTGGC	CATGAAGGTT	[664]
49_Strombomonas_triquetra	GCTCCGAAG	TGGACTTGAA	GTTCAGAGT	GCTGGCTCAT	ACCTAGATCC	CACAAAGGAT	GTTCGCTGCT	AGTGATAGCA	GGACGGTGGC	CATGAAGGTT	[695]
50_Eutreptiella_pomquetensis	GATCCGAGAG	CGTGCTTACA	GGGCCCGAGT	GCTGGCTCAT	GCCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGGCGGTGGC	CATGAAGGTT	[698]
51_Trachelomonas_planctonica	GCTCCGAGAG	TGGACTTACA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGACGGTGGC	CATGAAGGTT	[695]
52_Trachelomonas_lefevrei	GCTCCGAGAG	TGGACTTACA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGGGATAGCA	GGACGGTGGC	CATGAAGGTT	[697]
53_Trachelomonas_volvocina	GCTCCGAGAG	TGGGCTTACA	GATCCGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGAGATAGCA	GGACGGTGGC	CATGAAGGTT	[698]
54_Trachelomonas_conradi	GCTCCGAGAG	TGGACTTGCA	GATCCAGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGTGATAGCA	GGACGGTGGC	CATGAAGGTT	[681]
55_Trachelomonas_similis	GCTCCGAGAG	TGAGCTTACA	GATCCGAGT	GCTGGCTCAT	ACCCAGATCC	CACAAAGGAT	GTTCGCTGCT	AGAGATAGCA	GGACGGTGGC	CATGAAGGTT	[698]

	710	720	730	740	750	760	770	780	790	800	
01_Rhynchomonas_nasuta	GACATCCGCT	AAGGAGTGTG	T-----	-----	-----	-----	-----	-----	-----	-----	[447]
02_Blastocrithidia_culicis	GgCaTcCgCt	AaGgAgTgTg	TaAcAaCTCa	CcAaCcGaAt	CcAcCgG---	-----	-----	-----	-----	-----	[558]
03_Herpetomonas_muscarum	GGGATCCGCT	AAGGAGTGTG	TAACAA-CCA	CCAACCTGAAT	CCACCGGCCCC	-----	-----	-----	-----	-----	[547]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
05_Leishmania_infantum	GGCATCCGCT	AAGGAGTGTG	TAACAACTCA	CCAACCGAAT	CCACCGGCCC	CGAAAAATGGA	TGACGCTTAA	GCCAATTAGT	GATTGCCCAT	TATGTAGATC	[774]
06_Leptomonas_bifurcata	GGCATCCGCT	AAGGAGTGTG	TAACAACTCA	CCAACCGAAT	CCACCGG----	-----	-----	-----	-----	-----	[560]
07_Leishmania_major	GGCATCCGCT	AAGGAGTGTG	TAACAACTCA	CCAACCGAAT	CCACCGGCCC	CGAAAAATGGA	TGACGCTTAA	GCCAATTAGT	GATTGCCCAT	TATGTAGATC	[775]
08_Leishmania_tarentolae	GGCATCCGCT	AAGGAGTGTG	TAACAACTCA	CCAACCGAAT	CCACCGGCCC	CGAAAAATGGA	TGACGCTTAA	GCCAATTAGT	GATTGCCCAT	TATGTAGATC	[773]
09_Trypanosoma_cruzi	GGCATCCGCT	AAGGAGTGTG	TAACAACTCA	CCAACCGAAT	CCACCGGCCC	CGAAAAATGGA	TGACGCTCAA	GCCAATTAGT	GATTGCCCAT	TATGTAGATC	[773]
10_Trypanosoma_brucei	GGCATCCGCT	AAGGAGTGTG	TAACAACTCA	CCAACCGAAT	CCACCGGCCC	CGAAAAATGGA	TGACGCTTAA	GCCCCATGTG	CGTGATTGCC	CATGTAGACC	[774]
11_Bodo_saltans	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	GGAAATCCGCT	AAGGAGTGTG	TAACAACTCA	-----	-----	-----	-----	-----	-----	-----	[701]
15_Colacium_elongatum	GGAATCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	CGGTGCCCAA	GTCAAGTGAAG	CTGAAATTGC	CTTGTAGAAC	[798]
16_Cryptoglena_pigra	GGAACCCGCT	TAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CAAGCAGCTC	CGAAAAATGGA	TGAAGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTTGTAAAAAC	[798]
17_Cryptoglena_skujai	GGAACCCGCT	TAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CAAGCAGCTC	CGAAAAATGGA	TGAAGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTTGTAAAAAC	[784]
18_Colacium_sp	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTTGTAGAAC	[799]
19_Colacium_vesiculosum	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTTGTAGAAC	[800]
20_Euglena_clara	GGAACCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGCGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CATGATACCC	TTAGTAGAGC	[523]
21_Discoplastis_adunca	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CAGTCAGCCC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAGTGTGC	CTGTAGAAC	[796]
22_Euglena_deses	GGAAYCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAATGCCT	CACAGTCCAC	CGTGTAGAAC	[797]
23_Euglena_geniculata	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTTCCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CATAGTCCAC	CATGTAGAAC	[798]
24_Euglena_granulata	GGAACCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GCCAGCGAGA	CATACC-CTC	TGAGTAGAAC	[786]
25_Euglena_hiemalis	GGAACCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGCGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTCACA-GTC	TGCGTAGAAC	[794]
26_Eutreptia_pertyi	GGAATCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGAACAGCTC	CGAAAAATGGA	TGGTGCCCAA	GCCAGCGAAC	GTTA-TGCC	TGAGTAGGCC	[784]
27_Euglena_stellata	GGAATCCGCT	GAGGAGTGTG	TAACAACTCA	CCTGCCGAAG	CGAGCAGCCC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CACAATCGGC	CGTGTAGAAC	[783]
28_Euglena_longa	GGAACCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGCGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTCACA-GTC	TGCGTAGAAC	[781]
29_Euglena_viridis	GGAATCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CTTTACGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTCCCTCACA	ATCGTAGAAC	[798]
30_Euglena_exilis	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGGGCGGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGCGATCGC	CTCGTAGAAC	[798]
31_Euglena_gracilis	GGAACCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGCGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	TACAGTCTGC	CATGTAGAAC	[798]
32_Lepocinclis_steinii	GGAACCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGTCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAACCTGC	CTTGTAGAAC	[781]
33_Lepocinclis_acus_var_major	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TAGGCTGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGATATTGC	CTTGTAGGTC	[781]
34_Lepocinclis_spirogyroides	GGAACCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGCGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAACGTCG	CTTGTAGAAC	[783]
35_Lepocinclis_triapteris	GGAATCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTTGTAGAAC	[784]
36_Lepocinclis_ovum	GGAACCCGCT	GAGGAATGTG	TAACAACTCA	CCTGCCGAAT	CGCGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CCGAGTTTGT	CTTGTAGAAC	[784]
37_Monomorphina_aenigmatica	GGAACCCGCT	CAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CAAGCTGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	TTGAAGTTGC	CTCGTAGGCC	[798]
38_Colacium_mucronatum	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTTGTAGAAC	[799]
39_Monomorphina_sp	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	TGAATTTGCC	TCAGTAGAAC	[798]
40_Peranema_trichophorum	GGAATCCGCT	TAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGGGCGTAA	CGTAACTGA	TGACGCTTAA	GCTTAGTACT	TATCGTAGAC	CACGTAAGGC	[774]
41_Phacus_orbicularis	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGTGCAGTTT	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAACCTGC	CTTGTAGAAC	[798]
42_Phacus_similis	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	TGACAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAACCTGC	CTTGTAGAAC	[775]
43_Monomorphina_pyrum	GGAATCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	TGAATTTGCC	TCAGTAGAAC	[785]
44_Strombomonas_napiformis	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGGTGAGCTT	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAGTTCGC	TTAGCCACAC	[796]
45_Strombomonas_borysteniensis	GGAACCCGCT	AAGGAGTGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTCGTAGAAC	[796]
46_Strombomonas_acuminata	GGAACCCGCT	AAGGAGTGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAGTTTGC	CTTGTAGAAC	[781]
47_Strombomonas_ovalis	GGAACCCGCT	AAGGAGTGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	TGAATTTGCC	TCAGTAGAAC	[796]
48_Trachelomonas_echinata	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CAAGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTAATGTCGC	CTTGTAAAAAC	[764]
49_Strombomonas_triquetra	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTTGTAGAAC	[795]
50_Eutreptiella_pomquetensis	GGAATCCGCT	GAGGAGTGTG	TAACAACTCA	CCTGCCGAAG	CGGGCAGCTC	CGAAAAATGGA	TGACGCCCAA	GCCAGCGCAC	CGCCCTCGCC	TCAGTAGGTC	[798]
51_Trachelomonas_planctonica	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGAGCCCAA	GTCAAGTGAAG	CTGAATTTGC	CTTGTAGAAC	[795]
52_Trachelomonas_lefevrei	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGAGCCCAA	GTCAAGTGAAG	TGAATTTGCC	TATGTAGAAC	[797]
53_Trachelomonas_volvocina	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	TAAATTCGTC	TCAGTAGAAC	[798]
54_Trachelomonas_conradi	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGAGCCCAA	GTCAAGTGAAG	CTGAGTTTGC	CTTGTAGAAC	[781]
55_Trachelomonas_similis	GGAACCCGCT	AAGGAATGTG	TAACAACTCA	CCTGCCGAAG	CGTGCAGCTC	CGAAAAATGGA	TGGTGCCCAA	GTCAAGTGAAG	TGAATTCGTC	TCAGTAGAAC	[798]

	810	820	830	840	850	860	870	880	890	900	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[447]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[558]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[547]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----GCA	GATCTTG6TT	GGCGTAGCAA	AGATCTAACG	GAGATATACT	CAACA---TA	[50]
05_Leishmania_infantum	CTGGAGGT-T	AGGACGAAGC	TGGCGTGAGC	TAAGACCGGC	CTCTAGTGCA	GATCTTG6TT	GGCGTAGCAA	AGATCTAACG	GAGATATACT	CAACA---TG	[870]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[560]
07_Leishmania_major	CTGGAGGT-T	AGGACGAAGC	TGGCGTGAGC	TAAGACCGGC	CTCTAGTGCA	GATCTTG6TT	GGCGTAGCAA	AGATCTAACG	GAGATATACT	CAACA---TG	[871]
08_Leishmania_tarentolae	CTGGAGGT-T	AGGACGAAGC	TGGCGTGAGC	TAAGACCGGC	CTCTAGTGCA	GATCTTG6TT	GGCGTAGCAA	AGATCTAACG	GAGATATACT	CAACA---TG	[869]
09_Trypanosoma_cruzi	GTGGAGGTA-	AGGACGAAGC	TGGCGTGAGC	TCAGACCAGC	CTCTAGTGCA	GATCTTG6TT	GGCGTAGCAA	AGATCTAACG	GAAACTATT	TAACA---TG	[869]
10_Trypanosoma_brucei	GTGGGGGTG-	AGGACGAAGC	TGGCGCGAGC	TCGACCCGC	CTCTAGTGCA	GATCTTG6TT	GGCGTAGCAA	AGATCTAACG	GAATATTTT	TCTTAA--TG	[872]
11_Bodo_saltans	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[701]
15_Colacium_elongatum	GTGGGAGATT	GCAGAGCCAG	GGAT-TGCTC	CTGACGTGTC	TTCTGGTGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	CAGA--CACC	[895]
16_Cryptoglena_pigra	ATGGAAGACT	GCAGAGCCTG	GGGTTTGAC	CTGGCGTATC	TTCTAGAGCA	GATCTTG6TT	GCAGTAGCAA	GTACTCCACC	AGGGGAAGAG	CAGTACAACC	[898]
17_Cryptoglena_sukjai	ATGGAAGACT	GCAGAGCCTG	GGGCTTGCGC	CTGGCGTATC	TTCTAGCGCA	GATCTTG6TT	GCAGTAGCAA	GTACTCCACC	AGGGGAAGAG	CAGTACAACC	[884]
18_Colacium_sp	GTGGGAGATT	GCAGAGCCGG	GGGATTGCTC	CCGACGTGTC	TTCTGGTGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	CAGA--CACC	[897]
19_Colacium_vesiculosum	ATGGGAGATT	GCAGTGCCAG	GGGATTGCTC	CCAGCGTGTC	TTCTGGTGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	CAGAACCAACC	[900]
20_Euglena_clara	TTTGAGATT	GCAGAGA-CCG	GGAGCGGATT	CCAGCGCGTC	TTCTGGCGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCATC	GGGTGAAGTG	TGGCCAAAAC	[622]
21_Discoploastis_adunca	ATGGAAGCT-	GCAGTGCCCT	GTGGTTATCC	TCAGCGTGCC	TTCTGGAGCA	GATCTTG6TT	GCAGTAGCAA	GTACTCCACC	GGGTGAAGAG	TGGCTTGACC	[895]
22_Euglena_deses	ATGGGAGATT	GCAGAGCCTG	GGGGTCACCC	CTGGCGTGTC	TTCCGGAGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	CAGTGTACCC	[897]
23_Euglena_geniculata	ACTGGGACT	CGGTAGCCAG	GGGGCCGCC	CCGGCGGTC	TCCAGGAGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCATC	GAGGGTCGGC	AGAAACGCTC	[898]
24_Euglena_granulata	TCGGGCGACC	GCCGTGCGGG	CCCCCGGTC	CCATTGCGTC	GCCCGGCGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	CGGCCAAAAC	[886]
25_Euglena_hiemalis	ATGGATGACC	GCTGATCGTG	GGGAGTGTC	CCGGCGGTC	ACCCAGAGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCACC	GAATGAGGAG	TGGCAACATC	[894]
26_Eutreptia_pertyi	ACAGGGGCTT	GTGGTGCCGT	CGGGGCAACC	CGAATGTGCC	CTCTGGTGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCT--C	TCAGGTTGCG	TGCCGCAATT	[882]
27_Euglena_stellata	ATGAGGGACC	GCTGCTCCGG	GGGTCTGTCC	CCAGCGCGTC	CCCTAGAGCA	GATCTTG6TT	GTAGTAGCAA	GTACTCCATC	GAGTGAAGGG	CGGCAAACTC	[883]
28_Euglena_longa	ATGGATGACC	GCTGATCGTG	GGGAGTGTC	CCGGCGGCT	CCCCAGAGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCACC	GAATGAGGAG	TGGCAACATC	[881]
29_Euglena_viridis	ATGAGAGATT	GCAGAGCTGG	CAGGGCCAAAC	CTGGCGTGTC	TTCTGGTGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCATC	GGGTGAAGAG	TGGA-TAACC	[897]
30_Euglena_exilis	ACGGGAGCTT	GCCGTGCCCG	GGGTCAACCC	CCGGCGTGCC	TCCTGGCGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCATC	TGGAGAGGAG	CGTTTGGGCC	[898]
31_Euglena_gracilis	ATGGACAACA	GCTGATCGTG	GGGATTGATC	CCGGCGGCTT	GCCCAGAGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCACC	GAATGAGGAG	TGGCAACATC	[894]
32_Lepocinclis_steinii	TTGGGAGACT	GCTGTGCAAG	GGGGTTGACC	CTGGCGTGTC	TTCTGGCGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GAGTGAAGAG	TGGAACAACC	[881]
33_Lepocinclis_acus_var_major	ATGGGAGACC	GCAGTGCCCG	GGGACCGATC	CCAACGTGTC	TTCTGGCGCA	GATCTTG6TT	GCAGTAGCAA	GTACTCCATC	AGATGAAGAG	TGGAACAACC	[881]
34_Lepocinclis_spirogyroides	ACGGGAGGCT	GCCGTGCATG	GGGGTTGCC	CTGGCGTGCC	TTCTGGCGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	ATGTGAAGAG	TGGGAAAACC	[883]
35_Lepocinclis_triapteris	ATGGGAGATT	GCAGTGCCCT	GGGGCCAACC	CCAACGCGTC	TTCTGGCGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCATC	AGGTGAAGAG	TGGAACAACC	[884]
36_Lepocinclis_ovum	ATTGAGGACT	GCTGTGCAAG	GGGGGTGACC	CTGGCGGTC	CTCTGGCGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	CAGATCAACC	[884]
37_Monomorphina_aenigmatica	ATGGGAGATT	GCAGCGCCCG	GGGGTTGTCC	CCGGCGTGTC	TTCCAGCGCA	GATCTTG6TT	GCAGTAGCAA	GTACTCCATC	GGGTGAAGCG	CAGAACCAACC	[898]
38_Colacium_mucronatum	GTGGGAGACT	GCAGAGCCAG	GGGATTGCTC	CCAGCGTGTC	TTCTGGTGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	CAGAACCAACC	[899]
39_Monomorphina_sp	ATGGGAGATT	GCAGTGCCCT	GGGGTCAACC	CTGGCGTGTC	TTCTGGCGCA	GATCTTG6TT	GCAGTAGCAA	GTACTCCATC	GG-GGAAGT	GCAGAACACC	[897]
40_Peranema_trichophorum	ATGGTCGAG-	ACGTAGAAGA	AAGGTTGACC	TAAGTGCTC	GACCACTGCA	GATCTTG6TT	GCAGTAGCGG	ATA--CT--T	AG-T-----	-----CA	[854]
41_Phacus_orbicularis	TTGGGAGACG	GCGGCGCTGG	GGGGCCAACC	CCGGCGCACC	TCCCGGCGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCATC	GAGACAGGGG	CGGAACAACC	[898]
42_Phacus_similis	ATGGGAGCTT	GCAGTGAATG	AGGGTTGACC	CTGGCGTGCC	TTCTGGAGCA	GATCTTG6TT	GTAGTAGCGA	GTACTCCATC	AGGTGAAGGG	CAGAACCAACC	[875]
43_Monomorphina_pyrum	ATGGGAGATT	GCAGTGCCCG	GGGGTTGCC	CTGGCGTGTC	TTCTGGCGCA	GATCTTG6TT	GCAGTAGCAA	GTACTCCATC	GGGTGAAGTG	CAGAACCAACC	[885]
44_Strombomonas_napiformis	AGTTATCCCT	GTGGTA-ATT	CTGA-CACCT	CTGGCGTATC	CTGAACCTGT	GCACTTAACC	AAAGGATCGC	AA--GGCCAC	GC-T-----	-----TG	[877]
45_Strombomonas_borysteniensis	ATGGGAGACA	GCAGTGCCAG	GGGGTTGCC	CCAACGTGTC	TTCTGGAGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	TGGCCATAAC	[896]
46_Strombomonas_acuminata	ATGGGAGACA	GCTAAGCCCA	GGGGTTGACC	CTAGTGAGCA	CTAGTGAGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	TGGCCAAAAC	[881]
47_Strombomonas_ovalis	ATGGGAGACA	GCAGTGCCAG	GGGGTTGCC	CCAGCGTGTC	TTCTGGAGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	TGGCCATAAC	[896]
48_Trachelomonas_echinata	ATTGTGGACT	GCTGAAATG	AGGCTTGAGC	CTGGCGTGTC	TTCTGGCGCA	GATCTTG6TT	GCAGTAGCGA	ATACTCCATC	TGATGAAGAG	CAGATCAACC	[864]
49_Strombomonas_triquetra	ATGGGAGACA	GCAGTGCCAG	GAGGTTGACC	TCAGCGTGTC	TTCTGGCGCA	GATCTTG6TT	GCAGTAGCGA	GTACTCCATC	GGGTGAAGAG	TGGCAACACC	[895]
50_Eutreptiella_pomquetensis	GCGGGGGCCC	GCGGAGCCGG	TGGGGTGACC	CAGGCGCGCC	CCCCGGCGCA	GATCTTG6TT	GTAGTAGCGA	AT--ACT--C	A-ATCGGGGT	GCGAGCGGCC	[893]
51_Trachelomonas_planctonica	ATGGGAGACT	GCAGCA-ACA	GGGGTTGCC	CTGACGTGTC	TTCTGGAGCA	GATCTTG6TT	GTAGTAGCGA	ATACTCCATC	GGGTGAAGAG	CAGTACAACC	[894]
52_Trachelomonas_lefevrei	ATGGGAGACT	GCAGCA-ACA	GGGGTTGCC	CTGACGTGTC	TTCTGGTGCA	GATCTTG6TT	GTAGTAGCGA	ATACTCCATC	GGGTGAAGAG	CAGAAATAACC	[896]
53_Trachelomonas_volvocina	ATGGGAGATT	GCAGAGCAAG	GGGTTTGAAC	CTCAGTGTC	TTCTGGCGCA	GATCTTG6TT	GTAGTAGCGA	ATACTCCATC	GA-G-----	-----TG	[883]
54_Trachelomonas_conradi	ATGGGAGATT	GCAGTGCCAG	GGGC-TGCC	CCAGCGTGTC	TTCTGGCGCA	GATCTTG6TT	GTAGTAGCGA	ATACTCCATC	GGGTGAAGAG	CAGACACACC	[880]
55_Trachelomonas_similis	ATGGAAGATC	ACAGAGCAAG	GGGTTTGAAAC	CTCACGTATC	TTCTGGCGCA	GATCTTG6TT	GTAGTAGCGA	ATACTCCATC	GA-G-----	-----TG	[883]

	910	920	930	940	950	960	970	980	990	1000	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[447]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[558]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[547]
04_Leishmania_braziliensis	CAACGGAGC- --GGGG----	AAGGATTACT	CGTACACGAG	TTGTTCCGAG	TGAGCACAGT	TATGGCGGGT	TCGGAAAA--	-----	-----	-----	[121]
05_Leishmania_infantum	CAACGGAGC- --GGGG----	AAGGATTACT	CGTACACGAG	TTGTTCCGAG	TGAGCACAGT	TATGAAAAGG	GATGAAATTT	-TTAG----	ATATTT	-GGT	[957]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[560]
07_Leishmania_major	CAACGGAGC- --GGGG----	AAGGATTACT	CGTACACGAG	TTGTTCCGAG	TGAGCACAGT	TATTTATACT	TGCCAATCAT	--CTC----	TTGCTCTTTT	-----	[958]
08_Leishmania_tarentolae	CAACGGAGC- --GGGG----	AAGGATTACT	CGTACACGAG	TTGTTCCGAG	TGAGCACATT	ATAGAAAAGG	GATGAAATTT	-TTTT----	ATATAGTGGT	-----	[957]
09_Trypanosoma_cruzi	CAACGGAGC- --GGGG----	AAGGATTACT	CGTACACGAG	TTGTTCCGAG	TGAGCACAGT	TACGACAGTC	TTCTCTCTT	-CCAAAAAGT	GTGGGT	-GGA	[960]
10_Trypanosoma_brucei	CAACGGAGC- --GGGG----	AAGGATTACT	CGTACACGAG	TTGTTCCGAG	TGAGCACAGT	TACG--TT-A	TATACGCATT	-TTAT----	TAAACT	-GGA	[956]
11_Bodo_saltans	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[701]
15_Colacium_elongatum	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[995]
16_Cryptoglena_pigra	TGCTGAGGAC	TGAAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	ATGCAGGATG	CAGTGCCCCC	[997]
17_Cryptoglena_squajai	TGCTGAGGAC	TGAAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	ATGCAGGATG	CAGTGCCCCC	[983]
18_Colacium_sp	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[997]
19_Colacium_vesiculosum	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCTCCC	[1000]
20_Euglena_clara	CACTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACATG	CAAGACGGCC	C-CCTT---	GCTTCGAGTG6	[717]
21_Discoploplastis_adunca	ACTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TTAGCCGACG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGGG	AGCATGCAGG	CAGTGCCCCC	[995]
22_Euglena_deses	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGACG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	AGCGGGCCTG	TGGTGCCCCC	[996]
23_Euglena_geniculata	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGACG	TGAGCCAAGG	TAAAAACATG	CAAAACGGGCA	AGCAGGCTTG	CAGTGA-CCC	[997]
24_Euglena_granulata	CGCTGAGGAC	TGGAGTGGAG	CAGGCTTCGT	TGAATGTGGG	TCAGCCGACG	TGAGCCAAGG	TAAAAACACA	CAAAATGGACA	G-CAGGCTTG	CAGTGCCACC	[985]
25_Euglena_hiemalis	CGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGACG	TGAGCCAAGG	TAAAAACAAG	CAAAATGGGCT	-CCAGGCGTC	CAGTGCCCCC	[993]
26_Eutreptia_pertyi	CGAAGGTGG- --GACC	CGGCA TGCAGACCAT	TGAACGCGGG	TCAGTGCATG	CGAGCCAAGG	TAAGAACCA- T	CCAACCT-GAT	G-CTGGG-CG	TG-GG-CCCC	-----	[973]
27_Euglena_stellata	CGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGACG	TGAGCCAAGG	TAAAAACATG	CAAGTGGGCA	ATCAAGCTTG	CAGTGCCCCC	[983]
28_Euglena_longa	CGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGACG	TGAGCCAAGG	TAAGGAGCCA	GGCATGGGCT	-CCAGGCGTC	CAGTGCCCCC	[980]
29_Euglena_viridis	CACAGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGACG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	AGCTT----	GAGTGCCCCC	[993]
30_Euglena_exilis	CGCCGACGCT	GTGTGCGGTT	GAGGACTCGT	TGAATGTGGG	TCAGCCGACG	TGAGCCAAGG	TAAAAACAGA	CAAAATGGGCA	GGCCCGG--G	CAGCGCCCCC	[996]
31_Euglena_gracilis	CGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGACG	TGAGCCAAGG	TAAAAACAAG	CAAAATGGGCT	-CCAGGCATC	CAGTGCCCCC	[993]
32_Lepocinclis_steinii	CGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGACG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCT	--CAGGCATG	CAGTG-CCCT	[978]
33_Lepocinclis_acus_var_major	CGCTGAGGAC	TGGAGTGGAG	CAGGCTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACATG	CAATCGGGCG	AGCAGGCTTG	CAGTGCTCCC	[980]
34_Lepocinclis_spirogyroides	CGCTGAGGAC	TGGAGTGGAG	CAGGCTTCGT	TGAATGCGGG	TCAGTGCAGC	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[983]
35_Lepocinclis_triptyris	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAGTGCAGC	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[984]
36_Lepocinclis_ovum	TGCTGAGGAC	TGGAGTGGAG	CAGGCTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGCCAAGG	TAAAAACATG	CAAGATGGCA	AGCAGGCTTG	CAGTGCCCCC	[984]
37_Monomorphina_aenigmatica	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[997]
38_Colacium_mucronatum	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[999]
39_Monomorphina_sp	TGCCGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACAGG	CAAAATGGCA	AGCAGGCTTG	CAGTGCCCCC	[996]
40_Peranema_trichophorum	GGAACTGGAG	T-GGAG----	CAGGTTTCGT	CGCAAGGGGG	TTAGTCAACG	TAAGCCATGG	CAAGA-AC-T	TGCTTTAGT-	--GTG----	GACCA-CCG	[939]
41_Phacus_orbicularis	CGCTGAGGGC	TGGAGTGGAG	CAGGCTTCGT	TGAATGCGGG	TCAGTGCAGC	TGAGCC-ATG	GTA AAAACATT	CAAGTGGCA	AGCAGGCCAG	CAGTGCCCCC	[997]
42_Phacus_similis	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAACCGGCG	TGAGCCAAGG	TAAAAACATG	CAAAACGGGCA	AGCAGGCTTG	CAGTGCCCCC	[975]
43_Monomorphina_pyrum	TGCTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGTGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[984]
44_Strombomonas_napiformis	CACGGTAC-T	GAGAGCAGCA	GCATCAGATC	TGAGCATTGC	TTCAGGACCC	TGCGTTGTTG	TAAGTCGATG	CCAGAAGC--	-----	C ATCCGA-GGG	[964]
45_Strombomonas_borsteniensis	CACTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGCGG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[995]
46_Strombomonas_acuminata	CACTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TTAGCCGCGG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[980]
47_Strombomonas_ovalis	CACTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGCGG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[995]
48_Trachelomonas_echinata	TGCTGAGGAC	AGGAGTGGAG	CAGGCTTCGT	TGAATGTGGG	TCAGCCGACG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	G-AGCT---	TGAAGCTCGT	[960]
49_Strombomonas_triquetra	ACTTGAGGAC	TGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGCGG	TGAGCC-ATG	GTA AAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGACCCC	[994]
50_Eutreptiella_pomquetensis	CGCCGAGGAC	TGGAGTGGAG	TAGGTTTCGT	TGAACGCGGG	TCAGCCGATG	CGAGCCAAGG	CAATTGGCCC	CAAAACGGCA	ATGCGGG--G	GGACG-CCCC	[990]
51_Trachelomonas_planctonica	TGCTGAGGAC	AGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGTCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[994]
52_Trachelomonas_lefevrei	TGCTGAGGAC	AGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGACG	TGAGTCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[996]
53_Trachelomonas_volvocina	AAGAGCCGTT	GTGTGCGGTT	GAGGACACGT	TGAATGCGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACATG	CAAGATGGCA	AGCAGGCTTG	CAGTGCCCCC	[982]
54_Trachelomonas_conradi	TGCTGAGGAC	AGGAGTGGAG	CAGGTTTCGT	TGAATGCGGG	TCAGCCGCGG	TGAGCCAAGG	TAAAAACATG	CAAAATGGGCA	AGCAGGCTTG	CAGTGCCCCC	[980]
55_Trachelomonas_similis	AAGAGCCGCT	ATGTGCGGTT	GAGGACACGT	TGAATGCGGG	TCAGCCGACG	TGAGCC-ATG	GTA AAAACA-T	GAAGATGGCA	AGCAGGCTTG	CAGTGCCCCC	[981]

	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[447]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[558]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[547]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[121]
05_Leishmania_infantum	AAT-GCGAAA	CACTTGCCAG	GTAACAA-A-	---T-CAATC	CTCCCACGG-	TGAAATCCAC	G6AAAAACAA	GAAACTTTTC	CGTACTCAT-	CACCGCATAT	[1048]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[560]
07_Leishmania_major	TAC-TTCT-C	---G--CGTT	G-TTCG--G-	---A-GCGG-	GG--GC----	TCACTCCCTA	GTGTGTGTGC	GCTGGTCTTC	CGTACTCAT-	CACCGCATAT	[1035]
08_Leishmania_tarentolae	AAT-GCGAAA	CACTTGCCAG	GTAACAA-A-	---T-CAATC	CTCCCACGG-	TGAAATCCAC	GTTATTGGAA	A--ACAAAAC	CGTACTCAT-	CACCGCATAT	[1046]
09_Trypanosoma_cruzi	AAT-GCGAAA	CACTTGCCAG	GTGACCA-A-	---T-CAATC	CTCCCACGG-	TGAAATCCAC	TGCGCATTTAA	A--GCAAAAAC	CGTACTCAT-	CACCGCATAT	[1049]
10_Trypanosoma_brucei	AAT-GCGAAA	CACTTGCCAG	GTGACCC-A-	---T-CCATC	CTCCCACGG-	TGAAATCCAC	G6AAAAATGAA	A--ACAAAAC	CGTACTCAT-	CACCGCATAT	[1045]
11_Bodo_saltans	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[701]
15_Colacium_elongatum	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAGG	CTCATGCGGG	CTATCGAT-T	GCAAGCAAGG	TGGCTCCAAT	GAACTCCTGT	CGTACTGAAT	CACCGCAGCT	[1094]
16_Cryptoglena_pigra	TGAGGAGTTT	TCGTTTCCTC	CTGATGCAGG	CTTGAGTGGG	CTGTTGATTG	GCACCTCT-G	CAATGAACCT	C--CTGTGGT	CGTACTGAAT	CACCGCAGCA	[1094]
17_Cryptoglena_sujai	GAG-GAGTTG	TCGTTTCCTC	CTGATGCAGG	CTCATGTGGG	CTGTTGATTG	GCACCTCC-G	GGCTCCAATG	AACCTCCTGT	CGTACTGAAT	CACCGCAGCA	[1081]
18_Colacium_sp	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAG-	---G-CTCA-	TG--CGGGCT	ATCGGAGGGA	ATCGGTGGTC	ACTGCCAATT	CGTACTGAAT	CACCGCAGCT	[1089]
19_Colacium_vesiculosum	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAGA	CTCATGCGGG	CTATCGAT-T	GCAAGCCC-T	TGGCTCCAAT	GAAAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1098]
20_Euglena_clara	CTCCCTGAG	GAGTTCTCGT	TTCTCTCTGA	TGCAGACCCA	TGTGGGCTAT	GCACTCCG--	GGCTCCAATG	A-AAGCCCGA	CGAAATGCAG	CTAGGGAATC	[814]
21_Discoploplaxis_adunca	TGTGGAGTTT	TCGTTTCCTC	CTGATGCAGG	CTCATGCGGG	CTATTGATTG	ACACACCG-T	GGCTCCAAGT	AACAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1094]
22_Euglena_deses	GAG-GAGTTC	TCGTTTCCTC	CTGATGCAGA	CCCATGTGGG	CTATCGATTG	GCAAGCAAGG	GGYTCCAATG	A--A-CGGCT	CGTACTGTAT	CACCGCAGCT	[1092]
23_Euglena_geniculata	CCAAGGAGTT	GTCGTTTCCT	CTGATGCAGA	GATCCATGTG	GGCTATGGAT	TGTAGCAAGG	TGGCTCCAAT	GAAAGCCTGT	CGTACTGAAT	CACCGCAGCA	[1097]
24_Euglena_granulata	CCAGGAGTTC	TATCGTTTCC	TCCTGATGCA	GACCATATGT	GGCTATCGAT	TGGAGCAAGG	AACATGGCTG	TCACCTCCAT	CGTACTGCAT	CACCGCAGCA	[1085]
25_Euglena_hiemalis	GTG-GAGTTC	TCGTTTCCTC	CTGATGCAGA	CTCACGCGGG	CTATCGATTG	GCACCTTCGC	GGTTCCAATG	AACAGCCTGC	CGTACTGCAT	CACCGCAGCT	[1092]
26_Eutreptia_pertyi	AG-AGAGTTA	TCTTCTCCTT	CTGACACAAG	AA-A-GTGAG	CTTGCAATTG	ATGGGCAAGG	GACTTCACGG	AATGCCCAAT	CGTACCAGAT	CACCGCATTT	[1070]
27_Euglena_stellata	GAG-GAGTTG	TCGTTTCCTC	CTGATGCAAG	ACCCATGTGG	GCTATCGATC	GGCAGCAAGG	GG-CTCCAAT	GAAAGCTTGT	CGTACTGAAT	CACCGCAGCG	[1081]
28_Euglena_longa	GTG-GAGTTC	TCGTTTCCTC	CTGATGCAGA	CTCACGCGGG	CTATCGATTG	GCAAGCAAGG	GGTTCCAATG	AACAGCCTGC	CGTACTGCAT	CACCGCAGCT	[1079]
29_Euglena_viridis	GAG-GAGTTA	TCGTTTCCTC	CTGATGCAGA	CTCATGTGGG	CTATCGATTG	GCAAGCAAGG	TGGCTCCAAT	GAAAGCTTGT	CGTACTGGAT	CACCGCAGCA	[1092]
30_Euglena_exilis	CGGGGAGTTC	TCGTTTCCTC	CTGATGCAGG	GCTCATGCGG	GCTATCGATT	GGCAGCAAGG	GGCTCCAAGT	A--AGCCTGT	CGTACTGTAT	CACCGCAGCT	[1094]
31_Euglena_gracilis	GTG-GAGTTC	TCGTTTCCTC	CTGATGCAGA	CTCACGCGGG	CTATCGATTG	GCAAGCAAGG	GGTTCCAATG	AACAGCCTGC	CGTACTGCAT	CACCGCAGCT	[1092]
32_Lepocinclis_steinii	CCATGGAGTT	CTCGTTTCCT	CTGATGCAGG	TCTCATGTGG	GCTATCGATT	AGCAGCAAGG	TGGCTCCAAT	GAAAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1078]
33_Lepocinclis_acus_var_major	CGAGGAGTTC	TCGTTTCCTC	CTGATGCAGA	CACATGTGGG	CTATTGATTG	GCAAGCAAGG	TGGCTCCAAT	GA-AGCCTGT	CGTACTGAAT	CACCGCAGCT	[1079]
34_Lepocinclis_spirogyroides	CGTGGAGTTC	TCGTTTCCTC	CTGATGCAGG	GCTCACGTGG	GCTCTGGATT	GTTAGCAAGG	GGCTCCAAGT	A--AGCCTGT	CGTACTGAAT	CACCGCAGCT	[1081]
35_Lepocinclis_triapteris	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAGT	CTCATGTGGG	CTATGGATTG	GCAAGCAAGG	GGCTCCAAGT	AACAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1084]
36_Lepocinclis_ovum	TGGGGAGTTC	TCGTTTCCTC	CTGATGCAGT	TCCA-GTGGA	CTACCAATTG	TCAAGCAAGG	GGCTCCAAGT	AACAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1083]
37_Monomorphina_aenigmatica	TGAGGAGTTC	TCGTTTCCTT	CTGATGCAGG	TCCACGTGGG	CTATCGATTG	GCAAGCAAGG	GGCTCCAATG	A--AGTCTTT	CGTACTGAAT	CACCGCAGCT	[1095]
38_Colacium_mucronatum	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAGG	TTTCATGCGGG	CTATCG-A--	TTGGATAT-G	GGCTCCAATG	AACAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1095]
39_Monomorphina_sp	TGAGGAGTTA	TCGTTTCCTC	CTGATGCAGG	TTTCATGCGG	CTGTTGAATT	GCAGTCTG-T	CTTCTGGCAG	TACCTTCCAT	CGTACTGAAT	CACCGCAGCT	[1095]
40_Peranema_trichophorum	AGG-GTAA-T	---TCG-GTC	C-TGCT--A-	---A-TGCT-	GG--GG-A--	AGAAAGGGGG	CTAACCTAAC	A--CCCAGAT	TGTACCTAAG	ATCCACAAC	[1017]
41_Phacus_orbicularis	CGGGGAGTTC	TCGTTTCCTC	CTGATGCAGG	GTTTCATGTG	GCTATCGATC	GGCAGCAAGG	GGCCCCACTG	AACAGCCTGT	CGTACTGTAT	CACCGCAGCT	[1097]
42_Phacus_similis	TGAGGAGTTC	TCGTTTCCTC	CTGATGAAAG	TTTCATGTGG	CTATCGATTG	ACAAGCAAGG	GGCTCCAAGT	AACAGCCTGC	CGTACTGAAT	CACCGCAGCT	[1075]
43_Monomorphina_pyrum	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAGG	TTTCATGTGG	CTGTTGAATT	GCAAGCAAGG	AAATCTCTGG	C--AGTCACT	CGTACTGAAT	CACCGCAGCT	[1082]
44_Strombomonas_napiformis	ACAGGCTG-A	TGTCTAAGGG	GTCGGG--TT	CCCAGATCAC	TCAATG-A--	GCACACTGCT	TGCTACACCC	A--ACATGT	CAAACTCAAC	AGGTCTTCCC	[1055]
45_Strombomonas_borysteniensis	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAGG	TTCAAGTGGG	CTATGGATTG	CATTGCCCTC	GGCTCCAATG	A-AAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1094]
46_Strombomonas_acuminata	TGAGGAGTTA	TCGTTTCCTC	CTGATGCAGG	GCTCAAGTGG	GCTATGGATT	CAAGCAAGG	GGCTCCAATG	A-AAGCCTGT	CGTACTGCAT	CACCGCAGCT	[1079]
47_Strombomonas_ovalis	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAG-	---G-TTCAA	GTG-GGCT--	ATGGGTGG--	GGCTCCAATG	A-AAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1084]
48_Trachelomonas_echinata	TTT-CTCT-T	---GATAGAT	TGAGCAAGG	TG-A-GCCAT	TGATTGGCA-	TGACACTG--	GCCTCCAAGT	AACATCCTGT	CGTACTGAAT	CACCGCAGCA	[1050]
49_Strombomonas_triquetra	TGAGGAGTTC	TCGTTTCCTC	CTGATGCAGG	TTCAAGTGGG	CTATGGAT-T	GCAATGCCCC	TGGCTCCAAT	GAAAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1093]
50_Eutreptiella_pomquetensis	CGAGGAGTTG	TCTTTTCCTT	CTGACGCAAG	AGCACGCGGG	CTATCGATTG	GCAAGCAAGG	GGCTCCAAGG	A--AGCCCC	CGTACCAGAT	CACCGCAGCT	[1088]
51_Trachelomonas_planctonica	CGAGGAGTTC	TCGTTTCCTC	CTGATGCAGG	CACATGTGGG	CTATCGATTG	GCAAGCAAGG	TGGCTCCAAG	GAAAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1094]
52_Trachelomonas_lefevrei	CGAGGAGTTA	TCGTTTCCTC	CTGATGCAGG	CACATGTGGG	CTATCGATTG	GCAAGCAAGG	TGGCTCCAAG	GAAAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1096]
53_Trachelomonas_volvocina	CGAGGAGTTA	TCGTTTCCTC	CTGATGCAGA	CCCATGTGGG	CTATCGATTG	GCAAGCAAGA	TGGCCCCAAT	GAAAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1082]
54_Trachelomonas_conradi	CGAGGAGTTA	TCGTTTCCTC	CTGATGCAGG	CTCATGTGGG	CTATCGATTG	GCAAGCAAGG	GGCTCCAAGG	A-AAGCCTGT	CGTACTGAAT	CACCGCAGCT	[1079]
55_Trachelomonas_similis	CAGGGAGATA	TCTTCTCCTC	CTGATGCAGA	ATCCATGTGA	GCTATCGATT	TGCAAGAG-T	GCCTCATGGA	A--AGCCTGT	CGTACTGAAT	CACCGCAGCT	[1078]

	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[447]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[558]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[547]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[121]
05_Leishmania_infantum	AGAAGCTCTT	AGAGAAA-CA	GGGAAGTCGG	CAAAATAGAT	CGGTAAGTTC	GCAAGAAAGCA	TTGGCTCGAA	GGCGGCG-AC	CCGG-CGTCT	CACCTCTGGCA	[1145]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[560]
07_Leishmania_major	AGAAGCTCTT	AGAGAAA-CA	GGGAAGTCGG	CAAAATAGAT	CGGTAAGTTC	GCAAGAAAGCA	TTGGCTCAGA	ACTG--CGAT	CCAACTGTAT	AATTA AAAACA	[1132]
08_Leishmania_tarentolae	AGAAGCTCTT	AGAGAAA-NA	GGGAAGTCGG	CAAAATAGAT	CGGTAAGTTC	GCAAGAAAGCA	TTGGCTCAGA	ACTG--CGAT	CCAACTGTAT	AATTA AAAACA	[1143]
09_Trypanosoma_cruzi	ACAAGCTCTT	AGAGAAA-CA	GGGAAGTCGG	CAAAATAGAT	CGGTAAGTTC	GCAAGAAAGCA	TTGGCTCAGA	ACTG--CGAT	CCAACTGTAT	AATTA AAAACA	[1146]
10_Trypanosoma_brucei	CGAAGCTCTT	AGAGAAA-CA	GGGAAGTCGG	CAAAATAGAT	CGGTAAGTTC	GCAAGAAAGCA	TTGGCTCAGA	ACTG--CGAT	CCAACTGTAT	AATTA AAAACA	[1142]
11_Bodo_saltans	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[701]
15_Colacium_elongatum	AGACTCGCTC	GGACTAAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1194]
16_Cryptoglena_pigra	AGAACCTCTC	AGACGAAAAA	GGGAAGTCGG	CAAACTGGCT	GTGTACCTTT	GGAAGATACA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1194]
17_Cryptoglena_skujai	AGAACCTCTC	AGACGAAAAA	GGGAAGTCGG	CAAACTGGCT	GTGTACCTTT	GGAAGATACA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1181]
18_Colacium_sp	AGACTCGCTC	GGACTAAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1189]
19_Colacium_vesiculosum	AGACTCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1198]
20_Euglena_clara	GGAAAGTCGG	AAGCCAAGTG	CGCGAGCAGG	GAAGAACAGC	CGCGGCCCTC	CGGTGTCCCG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[914]
21_Discoplastis_adunca	TGACCCGCTC	AGACTGAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1194]
22_Euglena_deses	AGACTCGCTC	GGATGAAAAA	GGGAAGTCGG	CAAACTGGCT	CGCGACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1192]
23_Euglena_geniculata	AGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCT	CGGTACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1197]
24_Euglena_granulata	TGAATCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCT	GTGCACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1185]
25_Euglena_hiemalis	AGACTCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	GTGCACCTTT	GGAAGATACG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1192]
26_Eutreptia_pertyi	GAATCCGCTC	GTTACAAAAA	GGGAAGTCGG	CAAACTGGCT	CGGTACCTTC	GGAAGATGCG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1170]
27_Euglena_stellata	AGAATCGCTC	AGACGAAAAA	GGGAAGTCGG	CAAACTGGCT	CGCGACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1181]
28_Euglena_longa	AGACTCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	GTGCACCTTT	GGAAGATACG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1179]
29_Euglena_viridis	AGAATCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCT	CGCGACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1192]
30_Euglena_exilis	AGACCCGCTC	GGAACGAAAA	GGGAAGTCGG	CAAACTAGCC	GTGTACCTTC	GGAAGATGCG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1194]
31_Euglena_gracilis	AGACTCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	GTGCACCTTT	GGAAGATACG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1192]
32_Lepocinclis_steinii	AAACCCGCTC	GGATTGAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1178]
33_Lepocinclis_acus_var_major	CACCCGCTC	GGATCGAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1179]
34_Lepocinclis_spirogyroides	AACCCGCTC	GGATTGAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1181]
35_Lepocinclis_triapteris	GAACCCGCTA	GGAATGAAAA	GGGAAGTCGG	CAAACTAGCC	GTGTACCTTC	GGAAGATGCG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1184]
36_Lepocinclis_ovum	GGACCCGCTC	AGATTGAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTC	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1183]
37_Monomorphina_aenigmatica	GGACCCGCTC	AGAAGGAAAA	GGGAAGTCGG	CAAACTGGCT	GTGTACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1195]
38_Colacium_mucronatum	AGACTCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTC	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1195]
39_Monomorphina_sp	AGACCCGCTC	AGATTGAAAA	GGGAAGTCGG	CAAACTAGCT	GTGTACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1195]
40_Peranema_trichophorum	AGTGTGATC	AGGACAA-AA	GGGAAGTCGG	CAAAATGGCT	GA-----	-----	-----	-----	-----	-----	[1058]
41_Phacus_orbicularis	TGACCCGCTC	GGATTGAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTC	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1197]
42_Phacus_similis	AGACCCGCTC	GGATTGAAAA	GGGAAGTCGG	CAAACTAGCT	GTGTACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1175]
43_Monomorphina_pyrum	AGACCCGCTC	AGACCGAAAA	GGGAAGTCGG	CAAACTAGCT	GTGTACCTTC	GGAAGATGCA	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1182]
44_Strombomonas_napiformis	GTTCTTGGT	TCACCGAAG	GGACAGCAGG	AATCTCGTAT	CCATTCTATC	CGGTCAATAA	TTGGTT-TCC	CCT-----	-----	-----	[1127]
45_Strombomonas_borsteniensis	TGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTCGCAGAT	CTGACTGTTT	AATTGAAACC	[1194]
46_Strombomonas_acuminata	TGACCCGCTC	AGACGAAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTCGCAGAT	CTGACTGTTT	AATTGAAACC	[1179]
47_Strombomonas_ovalis	TGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTCGCAGAT	CTGACTGTTT	AATTGAAACC	[1184]
48_Trachelomonas_echinata	CGAACCCGCTC	AGACTAAAAA	GGGAAGTCGG	CAAACTGGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACC	[1150]
49_Strombomonas_triquetra	TGACCCGCTC	AGACGAAAAA	GGGAAGTCGG	CAAACTAGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTCGCAGAT	CTGACTGTTT	AATTGAAACC	[1193]
50_Eutreptiella_pomquetensis	AGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	CGGTACCTTT	GGAAGATGCG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1188]
51_Trachelomonas_planctonica	AGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1194]
52_Trachelomonas_lefevrei	AGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1196]
53_Trachelomonas_volvocina	AGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1182]
54_Trachelomonas_conradi	AGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1179]
55_Trachelomonas_similis	AGACCCGCTC	GGACGAAAAA	GGGAAGTCGG	CAAACTGGCC	ATGTACCTTT	GGAAGATGTG	GTGGCTCGGA	ACTTGCAGAT	CTGACTGTTT	AATTGAAACA	[1178]

	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[447]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[558]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[547]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[121]
05_Leishmania_infantum	CGGGCCGGTC	TTGGCGG---	-TTCAC--GC	TCGCCCTGTG	CCGCGTCCCA	CAAAATTGCCA	ACTCAGAACT	G-----CT--	TACGGCGGGG	AATCCAACATG	[1232]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[560]
07_Leishmania_major	TAGGTTTGTG	ATGCATCCA-	AGTGGTGATA	TCGCAAACTG	ATTTCTGCCC	AGTGCTCTGA	ATGTCAACGT	GACGAGATTC	ACCGACGCGC	GGGTAACGCG	[1231]
08_Leishmania_tarentolae	TAGGTTTGTG	ATGCATCCA-	AGTGGTGATA	TCGCAAACTG	ATTTCTGCCC	AGTGCTCTGA	ATGTCAACGT	GACGAGATTC	ACCGACGCGC	GGGTAACGCG	[1242]
09_Trypanosoma_cruzi	TAGGTTTGTG	ATGCATC-CA	AGTGGTGATG	TCGCAAACTG	ATTTCTGCCC	AGTGCTCTGA	ATGTCAACGT	GACGAGATTC	ACCGACGCGC	GGGTAACGCG	[1245]
10_Trypanosoma_brucei	TAGGTTTGTG	ATGCATCCA-	AGTGGTGATG	TCGCAAACTG	ATTTCTGCCC	AGTGCTCTGA	ATGTCAACGT	GACGAGATTC	ACCGACGCGC	GGGTAACGCG	[1241]
11_Bodo_saltans	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[701]
15_Colacium_elongatum	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTG	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTTACTGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1294]
16_Cryptoglena_pigra	CAGCTTTGTG	AAGGGACCAG	CTGGCTGATC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1294]
17_Cryptoglena_squajai	CAGCTTTGTG	AAGGGACCAG	CTGGCTGATC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1281]
18_Colacium_sp	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTG	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTTACTGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1289]
19_Colacium_vesiculosum	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTG	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTTACTGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1298]
20_Euglena_clara	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAAAGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1014]
21_Discoplastis_adunca	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCACTGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1294]
22_Euglena_deses	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTG	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1292]
23_Euglena_geniculata	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTA	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1297]
24_Euglena_granulata	CAGCTTCGGG	ATGGGGTCCC	CGGGCTGGCC	CCCCGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATCC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1285]
25_Euglena_hiemalis	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1292]
26_Eutreptia_pertyi	CAGCTTTGTG	AGGGGACCAG	CTGGCTGGTG	TCACGAGGTG	ATTTCTGCCC	AGTGCTCTGA	TTGTCAACGC	GAAGAGATTC	TGCCAAGCGC	GGGTTAACGG	[1270]
27_Euglena_stellata	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTG	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATCC	CAAGAGATTG	AGGCAAGCGC	GGGTAACGCG	[1281]
28_Euglena_longa	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1279]
29_Euglena_viridis	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTG	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1292]
30_Euglena_exilis	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1294]
31_Euglena_gracilis	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCGTGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1292]
32_Lepocinclis_steinii	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTA	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGCA	GAAGAGATTC	ATGCAAGCGC	GGGTAACGCG	[1278]
33_Lepocinclis_acus_var_major	CAGCTTTGTG	AAGTGATCAG	CTCGCTGGTC	CCTTGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAAGGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1279]
34_Lepocinclis_spirogyroides	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAAGGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1281]
35_Lepocinclis_triptyris	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1284]
36_Lepocinclis_ovum	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTA	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAAGGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1283]
37_Monomorphina_aenigmatica	CAGCTTTGTG	AAGGGACCCA	AGGGCTGATT	CCACGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTTAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1295]
38_Colacium_mucronatum	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTG	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTTACTGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1295]
39_Monomorphina_sp	CAGCTTTGTG	AAGGGATCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAACGC	GAAGAGATTC	AGTGAAGCGC	GGGTAACGCG	[1295]
40_Peranema_trichophorum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[1058]
41_Phacus_orbicularis	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1297]
42_Phacus_similis	CAGCTTTGTG	ACGGGACCAG	CTGGCTGGTA	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	AAAGAGATTG	AGCCAAGCGC	GGGTAACGCG	[1275]
43_Monomorphina_pyrum	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAACGC	GAAGAGATTC	AGTGAAGCGC	GGGTAACGCG	[1282]
44_Strombomonas_napiformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[1127]
45_Strombomonas_borysteniensis	CAGCTTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1294]
46_Strombomonas_acuminata	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1279]
47_Strombomonas_ovalis	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1284]
48_Trachelomonas_echinata	CAGCTTTGTG	AAGGGTCCAG	CTGGCTGGTC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1250]
49_Strombomonas_triquetra	CAGCTTTGTG	AAGGGACCAG	CTGGCTGGTT	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1293]
50_Eutreptiella_pomquetensis	CAGCTTCGTG	AGGGGACCAG	CGGGCTGTTT	CCACGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAACGC	AAAGAGATTG	AGCCAAGCGC	GGGTAACGCG	[1288]
51_Trachelomonas_planctonica	CAGCTTTGTG	ATGGGTCCAG	CTGGCTGGTC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1294]
52_Trachelomonas_lefevrei	CAGCTTTGTG	ATGGGTCCAG	CTGGCTGGTC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1296]
53_Trachelomonas_volvocina	CAGCTTTGTG	AAGGGTCCAG	CTGGCTGGTC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1282]
54_Trachelomonas_conradi	CAGCTTTGTG	ACGGGTCCAG	CTGGCTGGTC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1279]
55_Trachelomonas_similis	CAGCTTTGTG	AAGGGATCAG	CTGGCTGGTC	CCATGAGGTG	ACTTCTGCCC	AGTGCTCTGA	TTGTCAATGC	GAAGAGATTC	AGCCAAGCGC	GGGTAACGCG	[1278]

	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[447]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[558]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[547]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[121]
05_Leishmania_infantum	TATAATTAAA	ACATAGGTTT	GTGATGCATC	CAAG---TTG	GGTTTAGC-G	TAGTCTTCGG	GC--TATGCA	CCGTCTAGGT	-----TTGGG	AGTAATTTTA	[1321]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[560]
07_Leishmania_major	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCTTCCAAT	TAGAGACGCG	CATGAATGGA	TTAATGAGAT	TCCTCTGTCT	CCGAGTTACT	[1331]
08_Leishmania_tarentolae	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCTTCCAAT	TAGAGACGCG	CATGAATGGA	TTAATGAGAT	TCCTCTGTCT	CCGAGTTACT	[1342]
09_Trypanosoma_cruzi	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCTTCCAAT	TAGAGACGCG	CATGAATGGA	TTAATGAGAT	TCCTCTGTCT	CCGAGTTACT	[1345]
10_Trypanosoma_brucei	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCTTCCAAT	TAGAGACGCG	CATGAATGGA	TTAATGAGAT	TCCTCTGTCT	CCGAGTTACT	[1341]
11_Bodo_saltans	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[701]
15_Colacium_elongatum	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1394]
16_Cryptoglena_pigra	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1394]
17_Cryptoglena_skujai	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1381]
18_Colacium_sp	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1389]
19_Colacium_vesiculosum	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1398]
20_Euglena_clara	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGCCGCT	[1114]
21_Discoplastis_adunca	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1394]
22_Euglena_deses	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGCCGCT	[1392]
23_Euglena_geniculata	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTCGCT	[1397]
24_Euglena_granulata	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGCCGCT	[1385]
25_Euglena_hiemalis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTCGCT	[1392]
26_Eutreptia_pertyi	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGCTACT	[1370]
27_Euglena_stellata	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTCGCT	[1381]
28_Euglena_longa	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTCGCT	[1379]
29_Euglena_viridis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTCGCT	[1392]
30_Euglena_exilis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTCGCT	[1394]
31_Euglena_gracilis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTCGCT	[1392]
32_Lepocinclis_steinii	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1378]
33_Lepocinclis_acus_var_major	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1379]
34_Lepocinclis_spirogyroides	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1381]
35_Lepocinclis_triptyris	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1384]
36_Lepocinclis_ovum	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1383]
37_Monomorphina_aenigmatica	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1395]
38_Colacium_mucronatum	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1395]
39_Monomorphina_sp	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1395]
40_Peranema_trichophorum	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[1058]
41_Phacus_orbicularis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTCGCT	[1397]
42_Phacus_similis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAATGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1375]
43_Monomorphina_pyrum	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGGTGCT	[1382]
44_Strombomonas_napiformis	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[1127]
45_Strombomonas_borysteniensis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1394]
46_Strombomonas_acuminata	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1379]
47_Strombomonas_ovalis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1384]
48_Trachelomonas_echinata	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCC	[1350]
49_Strombomonas_triquetra	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1393]
50_Eutreptiella_pomquetensis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGCCGCG	[1388]
51_Trachelomonas_planctonica	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCC	[1394]
52_Trachelomonas_lefevrei	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCC	[1396]
53_Trachelomonas_volvocina	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1382]
54_Trachelomonas_conradi	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCC	[1379]
55_Trachelomonas_similis	CGGGAGTAAC	TATGACTCTC	TTAAGGTAGC	CAAAATGCCTC	GTCAACCAAT	TATTGACGCG	CATGAATGGA	TCAACGAGAT	TCCTGCTGTC	CCTAGTTGCT	[1378]

	1410	1420	1430	1440	1450	
01_Rhynchomonas_nasuta	-----	-----	-----	-----	-----	[447]
02_Blastocrithidia_culicis	-----	-----	-----	-----	-----	[558]
03_Herpetomonas_muscarum	-----	-----	-----	-----	-----	[547]
04_Leishmania_braziliensis	-----	-----	-----	-----	-----	[121]
05_Leishmania_infantum	TGCGGTAAC	CCCGTCCCC	CGGGAAG--	GGTGGTGTGC	TTTTTCTG-	-C [1369]
06_Leptomonas_bifurcata	-----	-----	-----	-----	-----	[560]
07_Leishmania_major	ATCTAGCGAA	ACGACAGTCA	AGGGAACGGA	CTTGAAGGGC	TAAGCGGGGA	AA [1383]
08_Leishmania_tarentolae	ATCTAGCGAA	ACGACAGTCA	AGGGAACGGA	CTTGAAGGGC	TAAGCGGGGA	-- [1392]
09_Trypanosoma_cruzi	ATCTAGCGAA	ACGACAGTCA	AGGGAACGGA	CTTGAAGGGC	TAAGCGGGGA	AA [1397]
10_Trypanosoma_brucei	ATCTAGCGAC	ACGACAGTCA	AGGGAACGGA	CTTGAAGGGC	TAAGCGGGGA	AA [1393]
11_Bodo_saltans	-----	-----	-----	-----	-----	[588]
12_Dimastigella_trypaniformis	-----	-----	-----	-----	-----	[180]
13_Rhynchobodo_sp	-----	-----	-----	-----	-----	[180]
14_Ploeotia_oblonga	-----	-----	-----	-----	-----	[701]
15_Colacium_elongatum	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TTAGCGGGGA	AA [1446]
16_Cryptoglena_pigra	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1446]
17_Cryptoglena_skujai	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1433]
18_Colacium_sp	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAACGGGGA	AA [1441]
19_Colacium_vesiculosum	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGTGTG	TCAGCGGGGA	AA [1450]
20_Euglena_clara	GTCTGGTGAA	CCAACAACCA	AGGGAACGGG	CTTGGATN--	-----	-- [1152]
21_Discoplastis_adunca	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1446]
22_Euglena_deses	GTCTGGTGAA	CCAACAACCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1444]
23_Euglena_geniculata	GTCTGGTGAA	CCAACAACCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1449]
24_Euglena_granulata	GTCCGGTGAA	CCAACAGCCA	AGGGAACN--	-----	-----	-- [1413]
25_Euglena_hiemalis	GTCTGGTGAA	CCAACAACCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1444]
26_Eutreptia_pertyi	GTCCGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGCG	TCAGTGGGGA	AA [1422]
27_Euglena_stellata	GTCTGGTGAA	CCAACAACCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1433]
28_Euglena_longa	GTCTGGTGAA	CCAACAACCA	AGGAACCGGG	CTN-----	-----	-- [1412]
29_Euglena_viridis	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGAN--	-----	-- [1429]
30_Euglena_exilis	GTCCGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGTGTG	TCAGCGGGGA	AA [1446]
31_Euglena_gracilis	GTCTGGTGAA	CCAACAACCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1444]
32_Lepocinclis_steinii	GTCCGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGCGTG	TCAGCGGGGA	AA [1430]
33_Lepocinclis_acus_var_major	GTCCGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGCGTG	TCAN-----	-- [1423]
34_Lepocinclis_spirogyroides	GTCCGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGCGTG	TCAGCGGGGA	AA [1433]
35_Lepocinclis_tripteris	GTCCGGTGAA	CCTACAGCCA	AGGGAACGGG	CTTGGGCGTG	TCAGCGGGGA	AA [1436]
36_Lepocinclis_ovum	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGCGTG	TCAGCGGGGA	AA [1435]
37_Monomorphina_aenigmatica	GTCTGGTGAA	CCAACCGCCA	AGGGAACGGG	CTTGGGTGTG	TCAGCGGGGA	AA [1447]
38_Colacium_mucronatum	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGTGTG	TCAGCGGGGA	AA [1447]
39_Monomorphina_sp	GTCTGGTGAA	CCAACCGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1447]
40_Peranema_trichophorum	-----	-----	-----	-----	-----	[1058]
41_Phacus_orbicularis	GTCCGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGG	AA [1449]
42_Phacus_similis	GTCCGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGTGTG	TCAGCGGGGA	AA [1427]
43_Monomorphina_pyrum	GTCTGGTGAA	CCAACCGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1434]
44_Strombomonas_napiformis	-----	-----	-----	-----	-----	[1127]
45_Strombomonas_borsteniensis	GTCTGGTGAA	CCAACAGCCA	AGGGAAC--	-----	-----	-- [1421]
46_Strombomonas_acuminata	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGN--	-----	-- [1416]
47_Strombomonas_ovalis	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1436]
48_Trachelomonas_echinata	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1402]
49_Strombomonas_triquetra	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	T-----	-- [1434]
50_Eutreptiella_pomquetensis	ATCCGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGGTGTG	TCAGCGGGGA	AA [1440]
51_Trachelomonas_planctonica	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1446]
52_Trachelomonas_lefevrei	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTN-----	-----	-- [1430]
53_Trachelomonas_volvocina	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTN	-----	-- [1422]
54_Trachelomonas_conradi	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1431]
55_Trachelomonas_similis	GTCTGGTGAA	CCAACAGCCA	AGGGAACGGG	CTTGGATGTG	TCAGCGGGGA	AA [1430]

Figure 3.3. Figure 1 and its legend from Triemer et al. (2006), reprinted with permission from the primary author, the Journal of Phycology and the Phycological Society of America (PSA). “FIG. 1. Rooted Bayesian phylogenetic tree based on combined small subunit (SSU) and partial-large subunit (LSU) rDNA sequences. Numbers on the internodes represent the posterior probability (pp) for the node. Only pp values above 0.80 are shown. Bold letters (A–N) at the major nodes are to facilitate the discussion of the nodes. Each taxon name is followed by its culture collection origin. If sequences came from different sources, then the SSU source is given before the slash and the LSU source comes after the slash. Genus names appearing within [] indicate the former genus of the taxon. The tree was outgroup rooted posteriorly with *Eutreptia viridis* and *Eutreptiella* species. The ingroup taxa form three well-supported major clades: a *Phacus/Lepocinclis* clade, a *Discoplastis* clade, and a large multi-taxon clade spanning the tree from *Euglena* to *Cryptoglana*. Shading delineates the subclades. Scale bar represents number of substitutions/site.”

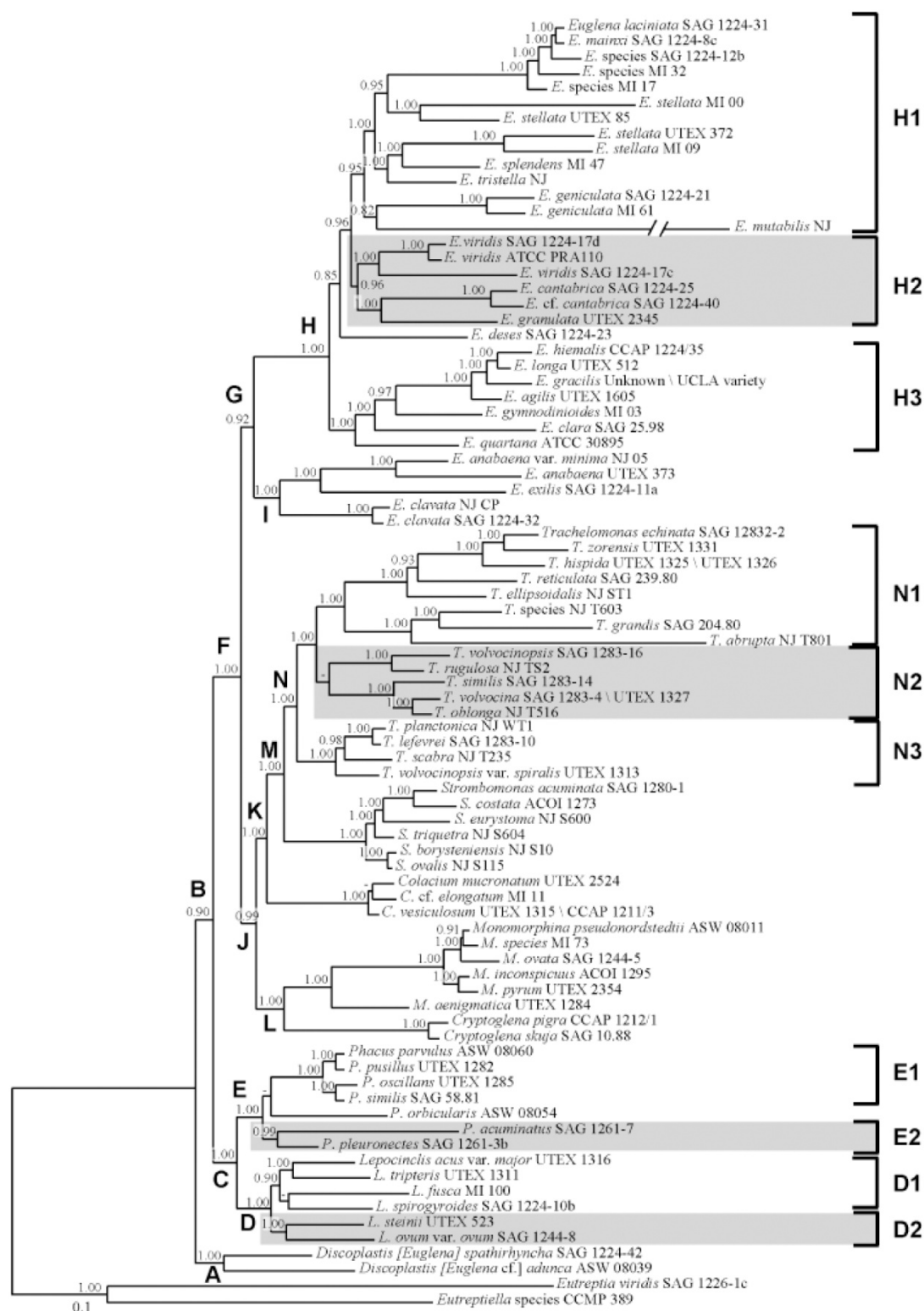


Figure 3.4. Cartoon topology summaries of single and concatenated gene ML and MLE analyses. Black=Euglenida, Het=Heteronematina (colorless euglenids), Eut=eutreptiales, D=*Discoplastis*, L=*Lepocinclis*, P=*Phacus*, E=*Euglena* (cI=clade I and cH=clade H from Triemer et al. (2006)), Cry=*Cryptoglana*, M=*Monomorphina*, Col=*Colacium*, T=*Trachelomonas*, S=*Strombomonas*, red and K=Kinetoplastida, the forced outgroup taxa in this study. End terminals of trees represent OTUs that are represented with support values of $\geq 50\%$ bootstrap (ML) or ≥ 0.9 posterior probability (MLE). Numbers in parentheses are the number of OTUs represented by the end terminal if less than the total number of species used in this study representing that genus or “group” compared to Figure 1 from Triemer et al. (2006). The OTUs that had support less than those values are not shown here and all other singletons are not shown. Asterisks after an abbreviated name or at a node represent that relationship of that OTU or clade was recovered in Figure 1 from Triemer et al. (2006). **A.** Summary of Figure 1 from Triemer et al. (2006), which used 18S concatenated with 28S, **B.** 1a, **C.** 1a3rd, **D.** 18S, **E.** 28S, **F.** 1a18S28S in black with changes to topology and clade recovery and support noted with blue dashes, an asterisk and an arrow for 1a3rd18S28S and the ML 18S28S analyses. The ML of 18S28S and both ML and MLE analyses of 1a3rd18S28S produced identical topologies and recovered clades, but overall 18S28S had decreased support at nodes compared to 1a3rd18S28S.

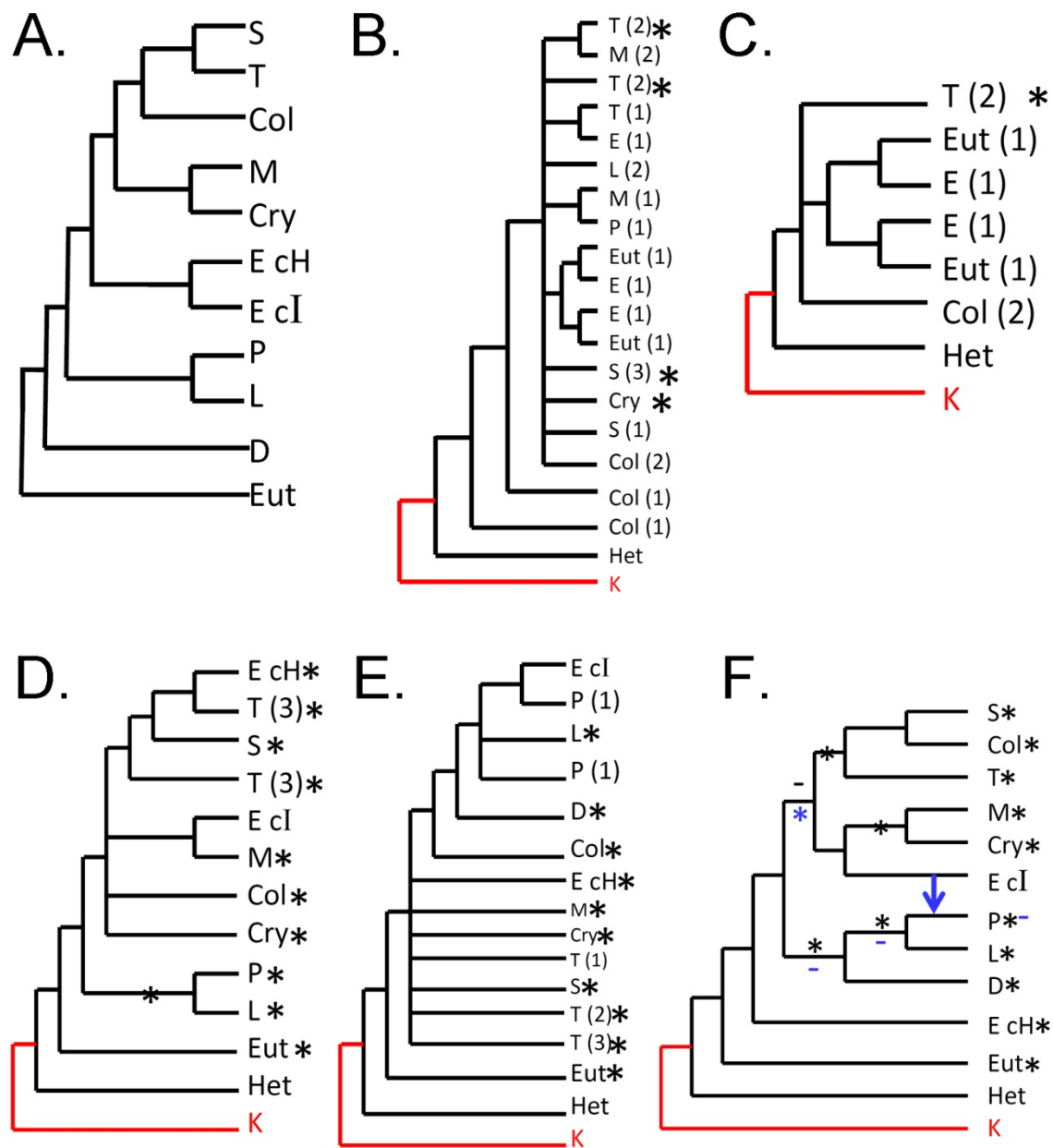


Figure 3.5. RAxML ML tree from the EF1 α single gene analysis (1a). The left or top value is the RAxML (ML) bootstrap score (bs) and the MrBayes (MLE) posterior probability (pp) is below or to the right. Only bs $\geq 50\%$ and pp ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bs and 1.00 pp recovered from both the ML and MLE analyses, respectively. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Genus and clade names are to the right even if they were not recovered as monophyletic. When a clade consisting of more than one OTU is recovered together as they were in Triemer et al. (2006), then the corresponding clade letter (and number if applicable) follows the clade name or vertical bar. If an identical relationship is recovered compared to Triemer et al. (2006), then the OTU(s) are shaded. Hash marks on a branch denote it is represented at one half the actual length.

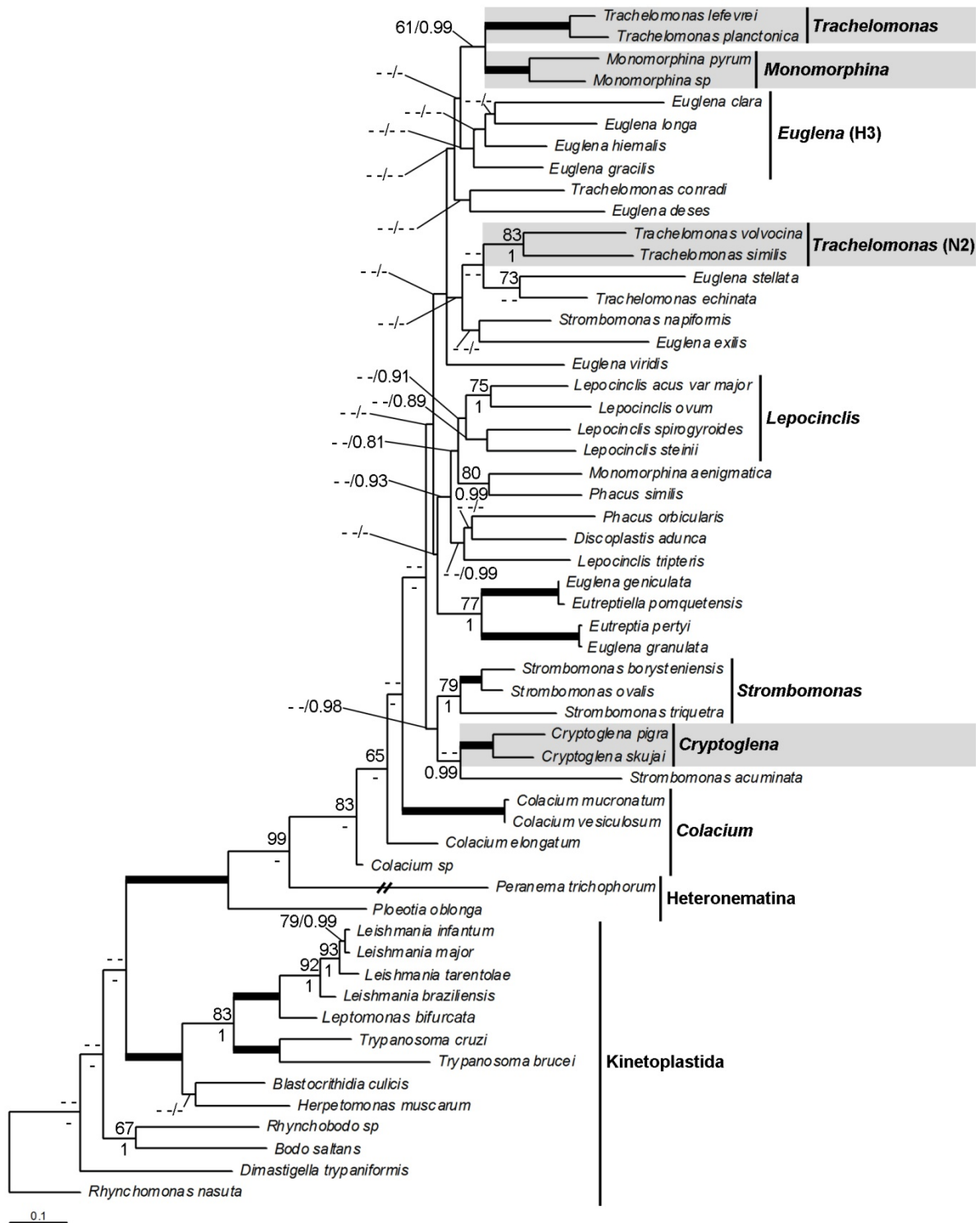


Figure 3.6. RAxML ML tree from the EF1 α single gene analysis with third codon positions removed (1a3rd). The left or top value is the RAxML (ML) bootstrap score (bs) and the MrBayes (MLE) posterior probability (pp) is below or to the right. Only bs $\geq 50\%$ and pp ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bs and 1.00 pp recovered from both the ML and MLE analyses, respectively. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Genus and clade names are to the right even if they were not recovered as monophyletic. When a clade consisting of more than one OTU is recovered together as they were in Triemer et al. (2006), then the corresponding clade letter (and number if applicable) follows the clade name or vertical bar. If an identical relationship is recovered compared to Triemer et al. (2006), then the OTU(s) are shaded. Hash marks on a branch denote it is represented at one half the actual length.

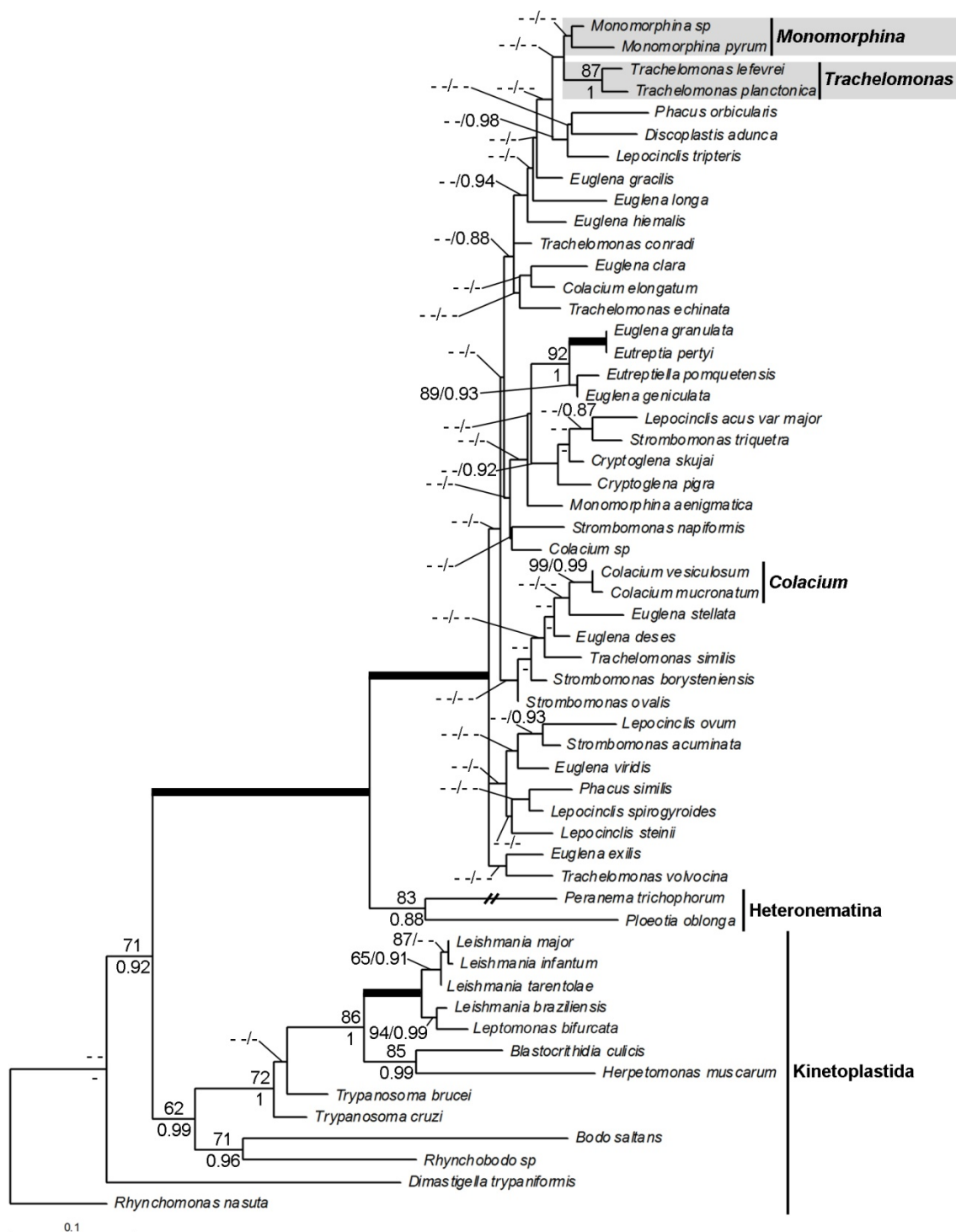


Figure 3.7. MrBayes consensus tree from the small subunit ribosomal DNA single gene analysis (18S). This tree was created using the GTR+ Γ MoSE with 10 million generations with 5 million generations used as burn-in. The left or top value is the RAxML (ML) bootstrap score (bs) and the MrBayes (MLE) posterior probability (pp) is below or to the right. Only bs $\geq 50\%$ and pp ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bs and 1.00 pp recovered from both the ML and MLE analyses, respectively. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Genus and clade names are to the right even if they were not recovered as monophyletic. When a clade consisting of more than one OTU is recovered together as they were in Triemer et al. (2006), then the corresponding clade letter (and number if applicable) follows the clade name or vertical bar. If an identical relationship is recovered compared to Triemer et al. (2006), then the OTU(s) are shaded. Hash marks on a branch denote it is represented at one half the actual length.

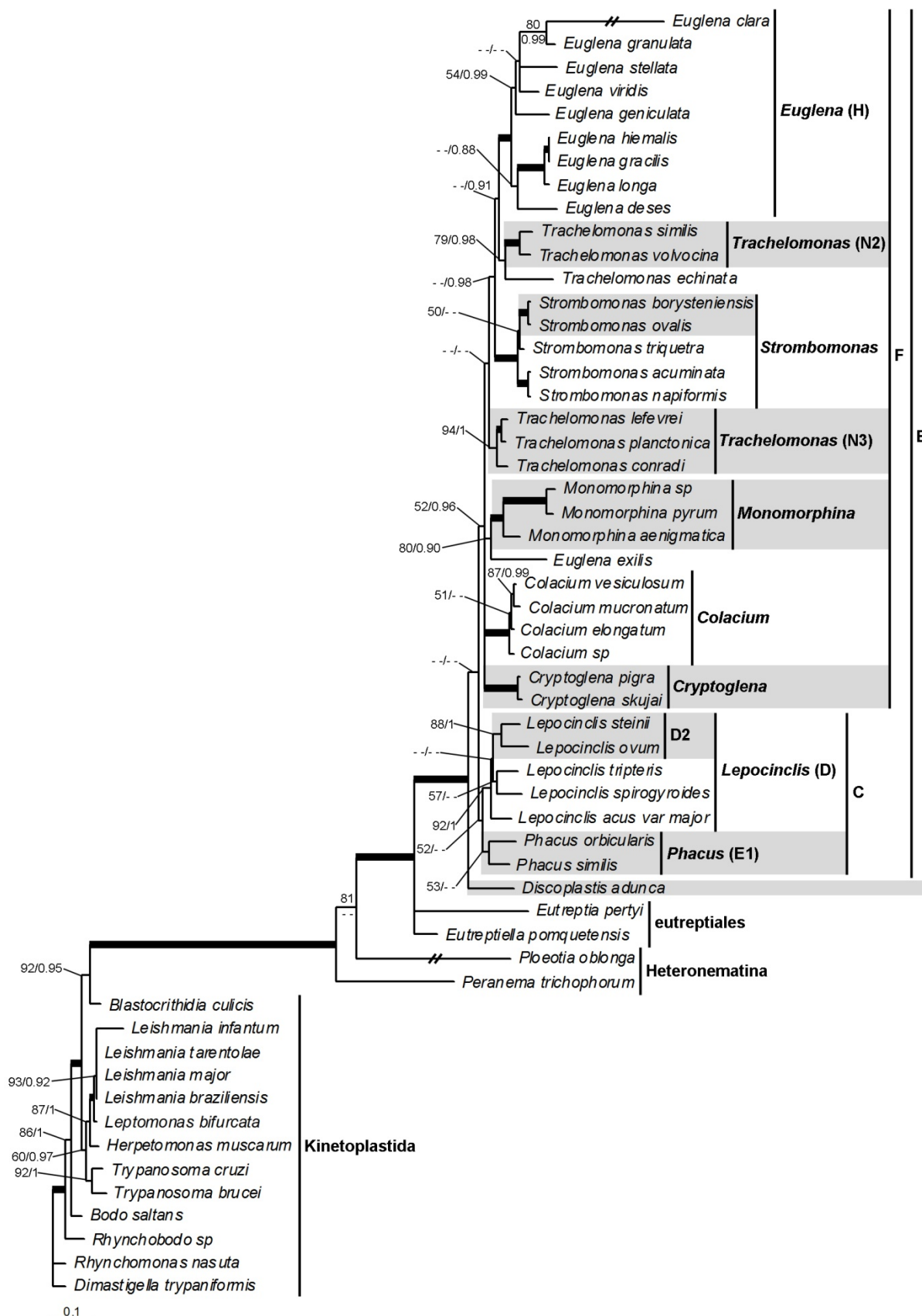


Figure 3.8. MrBayes consensus tree from the large subunit ribosomal DNA single gene analysis (28S). This tree was created using the GTR+ Γ MoSE with 10 million generations with 5 million generations used as burn-in. The left or top value is the RAxML (ML) bootstrap score (bs) and the MrBayes (MLE) posterior probability (pp) is below or to the right. Only bs $\geq 50\%$ and pp ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bs and 1.00 pp recovered from both the ML and MLE analyses, respectively. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Genus and clade names are to the right even if they were not recovered as monophyletic. When a clade consisting of more than one OTU is recovered together as they were in Triemer et al. (2006), then the corresponding clade letter (and number if applicable) follows the clade name or vertical bar. If an identical relationship is recovered compared to Triemer et al. (2006), then the OTU(s) are shaded. Hash marks on a branch denote it is represented at one half the actual length.

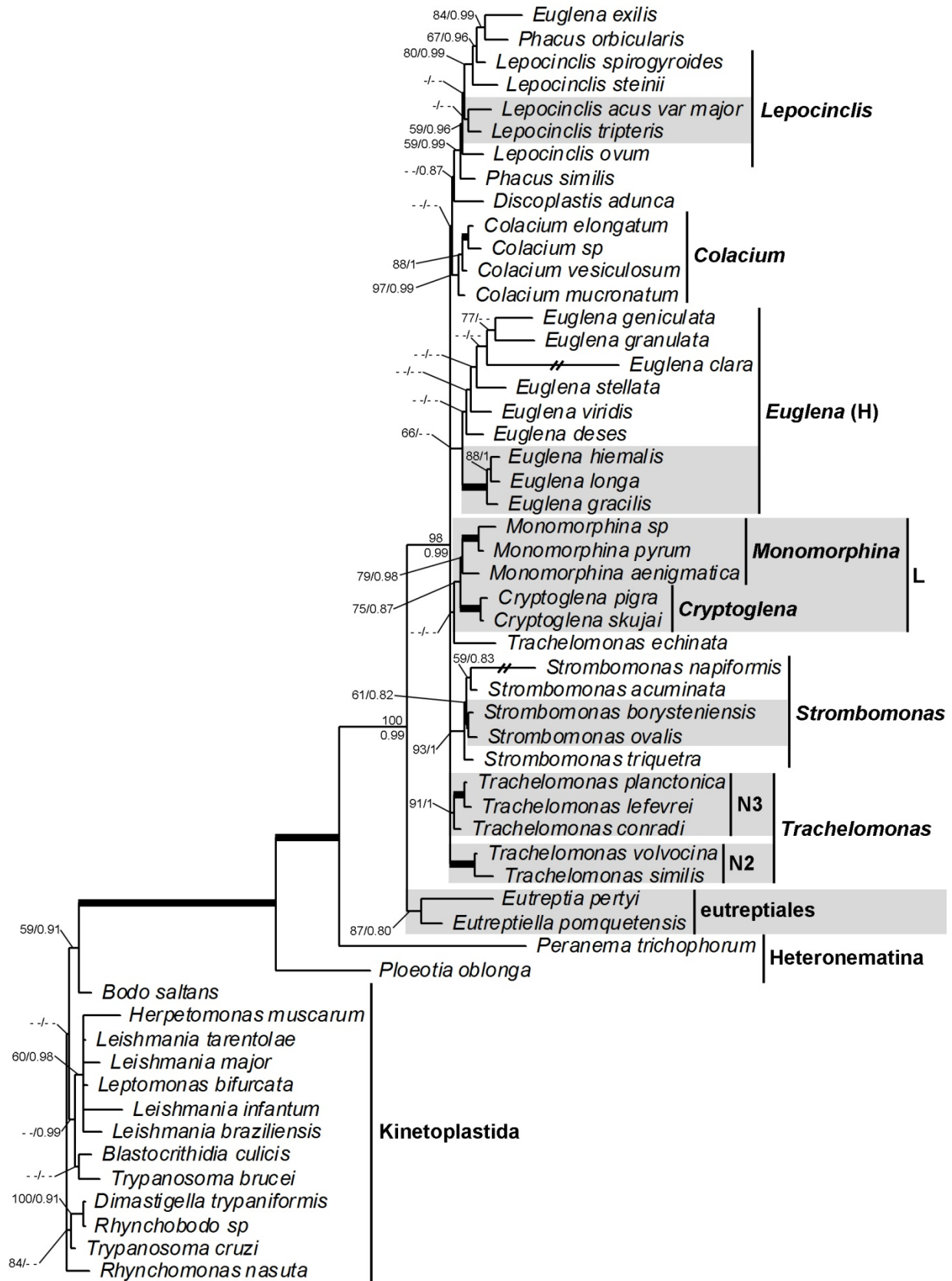


Figure 3.9. RAxML ML tree from the small and large subunit ribosomal DNA concatenated gene analysis (18S28S). This tree was created using the GTR+ Γ MoSE with 10 million generations with 5 million generations used as burn-in. The left or top value is the RAxML (ML) bootstrap score (bs) and the MrBayes (MLE) posterior probability (pp) is below or to the right. Only bs $\geq 50\%$ and pp ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bs and 1.00 pp recovered from both the ML and MLE analyses, respectively. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Genus and clade names are to the right even if they were not recovered as monophyletic. When a clade consisting of more than one OTU is recovered together as they were in Triemer et al. (2006), then the corresponding clade letter (and number if applicable) follows the clade name or vertical bar. If an identical relationship is recovered compared to Triemer et al. (2006), then the OTU(s) are shaded.

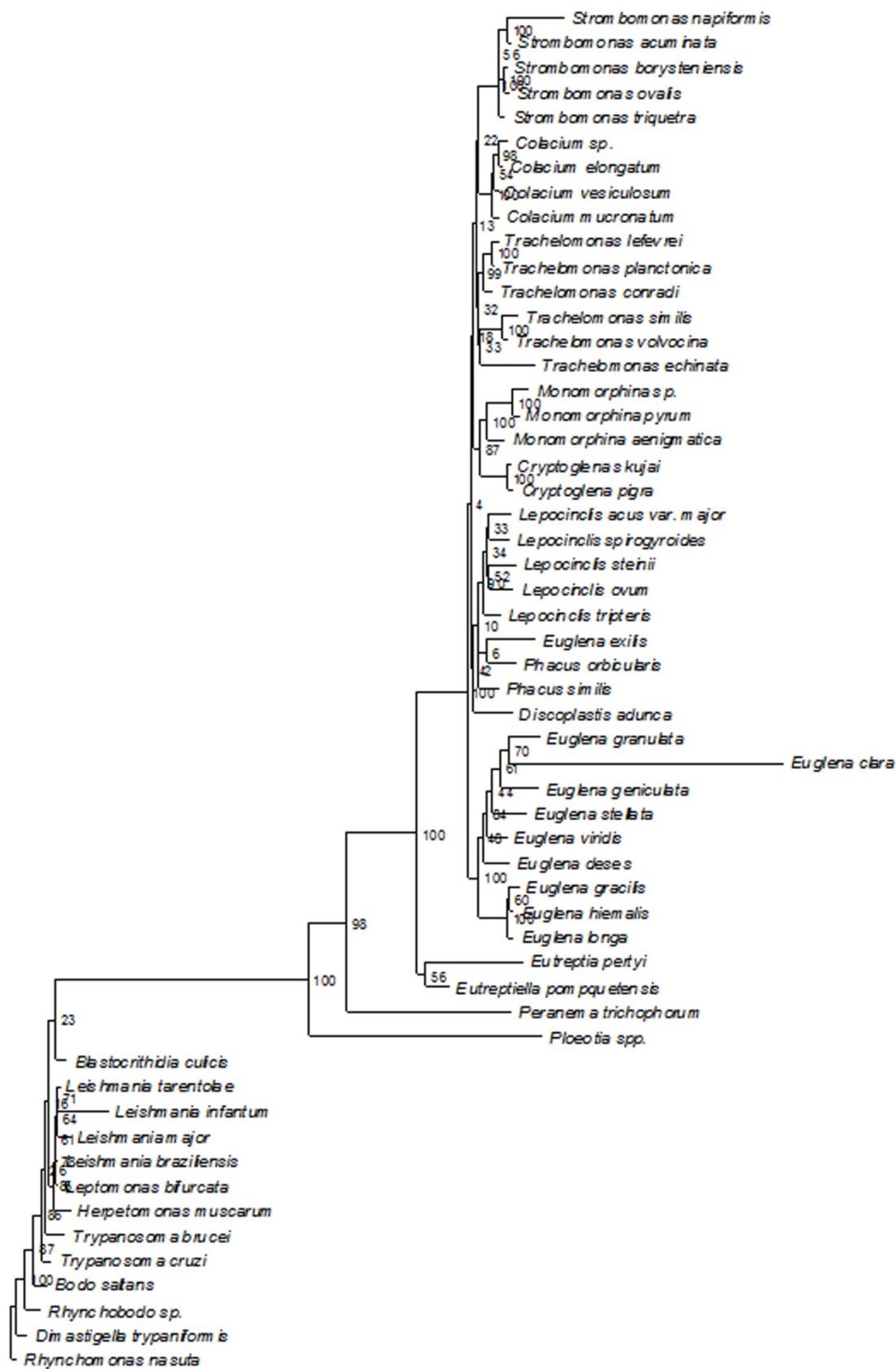


Figure 3.10. MrBayes consensus tree from the EF1 α , small and large subunit ribosomal DNA concatenated gene analysis (1a18S28S). This tree was created using the GTR+I+ Γ MoSE with 4 million generations with 2 million generations used as burn-in. The left or top value is the RAxML (ML) bootstrap score (bs) and the MrBayes (MLE) posterior probability (pp) is below or to the right. Only bs $\geq 50\%$ and pp ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bs and 1.00 pp recovered from both the ML and MLE analyses, respectively. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Genus and clade names are to the right even if they were not recovered as monophyletic. When a clade consisting of more than one OTU is recovered together as they were in Triemer et al. (2006), then the corresponding clade letter (and number if applicable) follows the clade name or vertical bar. If an identical relationship is recovered compared to Triemer et al. (2006), then the OTU(s) are shaded.

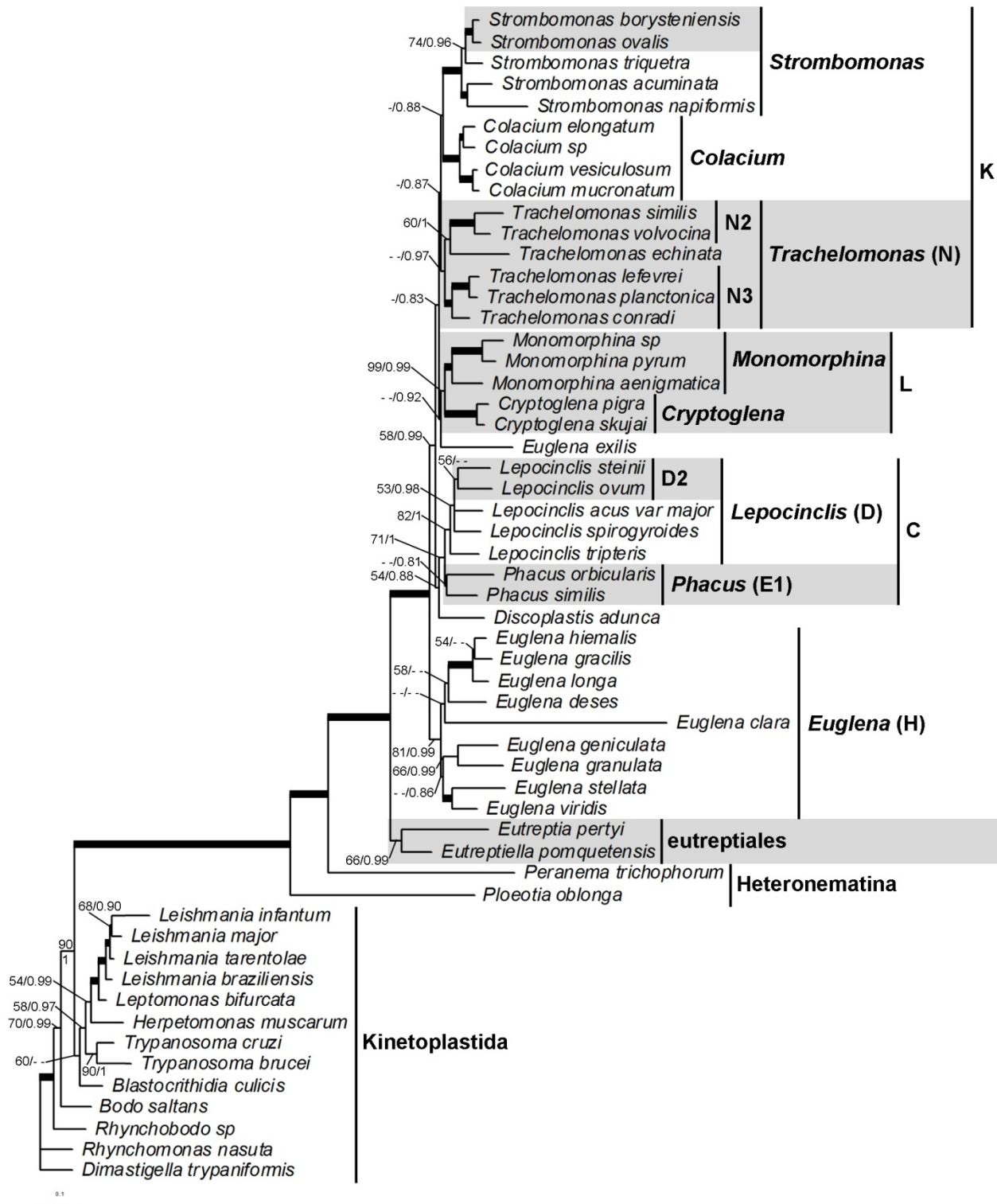
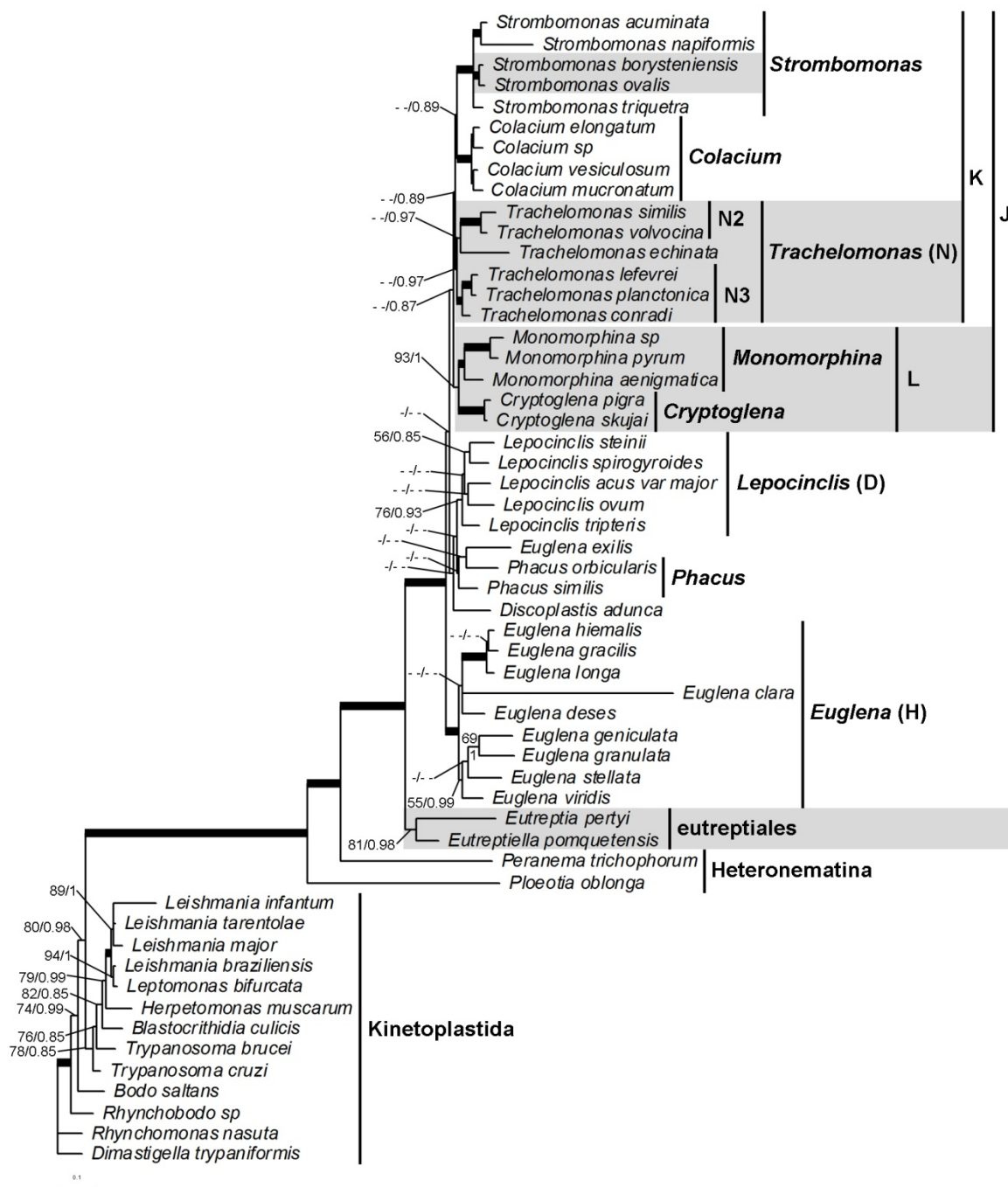


Figure 3.11. MrBayes consensus tree from the EF1 α with third codon positions removed, small and large subunit ribosomal DNA concatenated gene analysis (1a3rd18S28S). This tree was created using the GTR+I+ Γ MoSE with 4 million generations with 2 million generations used as burn-in. The left or top value is the RAxML (ML) bootstrap score (bs) and the MrBayes (MLE) posterior probability (pp) is below or to the right. Only bs $\geq 50\%$ and pp ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bs and 1.00 pp recovered from both the ML and MLE analyses, respectively. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Genus and clade names are to the right even if they were not recovered as monophyletic. When a clade consisting of more than one OTU is recovered together as they were in Triemer et al. (2006), then the corresponding clade letter (and number if applicable) follows the clade name or vertical bar. If an identical relationship is recovered compared to Triemer et al. (2006), then the OTU(s) are shaded.



CHAPTER 4

EF1 α EVOLUTION IN THE EUGLENID LINEAGE**Abstract:**

Euglenids (Euglenozoa/Excavata) are a diverse group of organisms that are represented by autotrophs, heterotrophs and osmotrophs. There has been a recent effort to incorporate protein-coding genes into uncovering the phylogenetic history of the group because those genes are supported as being more conserved and less prone to long branch attraction artifacts (LBA) as compared to the ribosomal genes. However, protein-coding genes can be members of complex gene families and have multiple copies of a gene in a lineage or organism due to independent gene duplication events. Elongation factor 1 alpha (EF1 α) has been used effectively to clarify relationships of organisms at the equivalent taxonomic level as that of the euglenids. It is a conserved, nuclear-encoded, protein-coding gene that is thought to be effectively single copy in eukaryotes and therefore is preferred as a phylogenetic marker. However, when recovering the single-gene phylogenies of euglenids with EF1 α there was evidence that supported the alternative hypotheses that either there are multiple gene copies of EF1 α in some of the lineages within the euglenids or that there is some type of editing of the genomic DNA (gDNA) before it is transcribed into mRNA. Here we show evidence of several independent gene duplication events in the Euglenida lineage and speculate on the evolution of EF1 α in the euglenid lineage as well as comparisons to one of the sister lineages of the euglenids, the kinetoplastids.

Introduction:

The Euglenida are ecologically important unicellular protists that are found within the earliest branching eukaryotic supergroup, Excavata (Xavier, Mainardespinto, Takino 1991; Baldauf 2003; Burki, Pawlowski 2006; Burki, Shalchian-Tabrizi, Pawlowski 2008). Previous clades and nomenclature of euglenids were based primarily on morphological structures that were visible with light microscopy (Leedale 1967; Triemer, Farmer 2007). Recent molecular evidence has revealed that clades and genera that were believed to be well supported were in fact not monophyletic and major revision of the group was required (Leander, Farmer 2001; Brosnan et al. 2003; Brown, Zakrys, Farmer 2003; von der Heyden et al. 2004; Triemer et al. 2006; Ciugulea et al. 2007; Ciugulea et al. 2008). Many of the currently described genera were shown to be polyphyletic and the taxonomy of the euglenids required a major revision.

Most studies of euglenid systematics have been based on analyses of sequences from ribosomal genes such as nuclear small subunit ribosomal DNA (18S), nuclear large subunit ribosomal DNA (28S) or plastid small or large subunit ribosomal DNA (Zakrys 1988; Linton et al. 2000; Brosnan et al. 2003; von der Heyden et al. 2004; Milanowski et al. 2006; Triemer et al. 2006; Ciugulea et al. 2007; Kim, Shin 2008; Kosmala et al. 2009). Allegedly, there is little phylogenetic signal left in at least the 18S molecular marker (Busse, Preisfeld 2002; Busse, Preisfeld 2003a). Similarly, 18S is prone to long branch attraction (LBA) artifacts leading to strongly supported yet incorrect relationships of taxa. The relationships of subgroups within the Euglenozoa could be robustly resolved if appropriate novel markers are used, especially when used in concatenation with other phylogenetic data to increase hidden underlying support (de Queiroz, Gatesy 2007). There has been a recent effort to incorporate protein-coding genes into uncovering the phylogenetic history of the group because those genes are supported as being

more conserved and less prone to LBA when compared to the ribosomal genes (Breglia, Slamovits, Leander 2007). However, protein-coding genes can be problematic because some of them have complex evolutionary histories including: members of complex gene families with multiple copies of genes, gene loss, or horizontal gene transfer (HGT) as reported in Chapter 1 and 2 (Simpson, Roger 2004). Previous chapters have discussed the utility of using the highly conserved, nuclear-encoded protein-coding gene elongation factor 1 alpha (EF1 α). This gene has been used effectively to determine relationships of organisms at the level of the earliest branching eukaryotes (Baldauf et al. 2000; Baldauf 2003) as well as the more recently diverged lineages that are similar to the level of the Euglenida (Baldauf, Doolittle 1997; Tanabe et al. 2004; Fiore-Donno et al. 2005). However, certain inconsistencies in the euglenid groups between the morphological and ribosomal lines of evidence (Leander, Farmer 2000; Leander, Witek, Farmer 2001; Brosnan et al. 2003; Brown, Zakrys, Farmer 2003; Brosnan et al. 2005; Triemer et al. 2006; Ciugulea et al. 2007; Ciugulea et al. 2008) versus the phylogeny created using EF1 α (Ch. 1, 2, 3) warranted a closer examination of the evolutionary history of EF1 α in the Euglenida.

The original publication and description of EF1 α in a euglenid was in the model organism, *Euglena gracilis* (Montandon, Stutz 1990). Montandon and Stutz (1990) produced a cDNA sequence of EF1 α that seemed to be functionally and phylogenetically equivalent to EF1 α in other distantly-related lineages. However, there was also evidence based on Southern blot analysis that there were at least four copies of EF1 α in the genome of *E. gracilis*. Montandon and Stutz (1990) posited that the other three copies were pseudogenes with either repetitive elements or introns that led them to have increased restriction sites as compared to the

published sequence, which they described as “the major stable transcript [of EF1 α that] is of the expected length”.

Organisms that are distantly-related to *E. gracilis* have been shown to have multiple copies of EF1 α in their genomes that are expressed either in different tissues or under different developmental stages (Vinkenoog et al. 1998; Mochida, Matsubara 2002): however, the sequence identity of the EF1 α copies in these organisms are such that the gene is still considered effectively a single-copy nuclear-encoded gene (Kidou, Ejiri 1998). Therefore, this was not considered problematic for phylogeny reconstruction of Euglenida using EF1 α . However, recent studies have also found lineage-specific gene duplications that are phylogenetically dissimilar from closely related organisms based on EF1 α (Akamine, Winter 2008). These findings coupled with the incongruent euglenid EF1 α phylogenies with previous morphological and ribosomal phylogenies led us to further analyze EF1 α from as many euglenids as possible to reconstruct the evolutionary history of EF1 α in the Euglenida.

Assuming that the life cycle stage of the organisms we harvested total genomic DNA (gDNA) and messenger RNA (mRNA) from were haploid, we were able to detect at least one, and possibly four, independent duplication events of EF1 α in the euglenids. These gene duplication events seem to have occurred after the divergence of the Euglenida and the Kinetoplastida based on evidence that the completed genomes of four trypanosomatids have effectively single copy versions of EF1 α . There seems to be evidence of the possible utility of EF1 α duplication events to track the phylogenetic history of some euglenid lineages due to the recent nature of the duplication events, such as that found in *Cryptoglena*. The timing of duplication events that have led to multiple divergent copies of EF1 α which do not reflect the organismal history of the group is unable to be resolved. Therefore, we conclude that EF1 α is

still phylogenetically informative, but that caution must be taken as to which group of organisms it is being used to study and which level of phylogeny reconstruction is being attempted.

Materials and Methods:

Culture information and nucleic acids extractions

All taxa with sources and accession numbers used in this study are shown in Table 4.1. Taxa were either obtained from culture collections or from colleagues (Triemer et al. 2006). All isolates were grown in the conditions and with the media specified in Chapter 1 of this thesis.

Sequence retrieval, nucleotide isolation, amplification, and sequencing

EF1 α sequences were acquired either with PCR or searches of EST and genomic databases (TBestDB, TriTrypDB v1.1) or GenBank release 172 (Table 4.1). New EF1 α sequences (17), representing 12 euglenid species were deposited in GenBank with accession numbers XXXX-XXXX. Total genomic DNA and/or total RNA was extracted from uniprotistan cultures and amplified and sequenced as specified in Chapter 1 of this thesis. We determined the copy numbers of EF1 α in the kinetoplastids that have had their entire genomes sequenced and mapped using OrthoMCL (Li, Stoeckert, Roos 2003) via the TriTrypDB v1.1: *Leishmania major*, *L. infantum*, *Trypanosoma brucei* and *T. cruzi* (Table 4.2). The *T. cruzi* genome is comprised of two different strains, and only those EF1 α paralogs from *T. cruzi* CL Brener Esmeraldo were used in this analysis, since this strain has a physical map. The actual location of each paralog was determined to be non-overlapping. They were also mapped onto the physical map of the haploid genomes of the kinetoplastid species analyzed such that alleles of a gene from a diploid genome are not confounding the genome copy number quantification. This allowed us to determine if the multiple distinct copies of EF1 α in certain euglenid lineages was specific to the ancestor of all euglenids or if they were possibly present in the ancestor of all Euglenozoa.

Sequence Alignment

All 60 euglenid nucleotide sequences of EF1 α , representing 42 species, were aligned with MUSCLE (Edgar 2004) and then edited by eye against closely related organisms' EF1 α sequences based on amino acid conceptual translations in Geneious Pro v3.8.5 (Biomatters Ltd.; available through www.geneious.com) (Drummond et al. 2009) and previously published EF1 α alignments (Baldauf, Palmer 1993; Baldauf, Doolittle 1997; Keeling, Inagaki 2004; Fiore-Donno et al. 2005; Gile, Patron, Keeling 2006; Gile, Keeling 2008). All resulting gapped alignments from EF1 α were used in their entirety to help provide enough tip-level signal to clarify the relationships of the different species of euglenids and the different gene duplicate clone types of EF1 α in some of those organisms. This resulted in an alignment with 1167 base pairs (bp) (Fig. 4.1). The proportion of gaps and completely undetermined characters in this alignment was 0.047229 as estimated by RAxML (see details of run below).

All hits from OrthoMCL in the TriTryp database 1.1 from the BLAST against *Leishmania braziliensis* EF1 α (TriTrypDB LBrM17-v2.0090) were downloaded from the website in a FASTA format. These were also run through MUSCLE and edited by eye based on the amino acid conceptual translations, but because they were nearly identical they were not compared to other organisms' EF1 α sequences. This gapped sequence alignment was 1352bp in length (Fig. 4.2)

Phylogenetic analysis

The euglenid alignment was analyzed with Maximum likelihood (ML), and Maximum likelihood estimation using Bayesian algorithms (MLE). ML was performed using RAxML v7.0.4 with simultaneous autodetect rapid bootstrapping (Stamatakis 2006; Stamatakis, Hoover, Rougemont 2008) on the CIPRES server (Miller et al. 2009). Two hundred and fifty bootstrap

replicates were performed for RAxML. MLE was also performed with MrBayes v3.1.2 (Huelsenbeck, Ronquist 2001; Ronquist, Huelsenbeck 2003) on the University of Georgia (UGA) rCluster. The MLE analysis parameters were as follows for each of the four identical runs: McMC parameters, nchains=4; Temp=0.1; Swapfreq=1; Nswaps=1, 4 gamma categories; empirical frequencies were estimated, 30 million generations sampled every 1000 generations with 15 million generations in the burn in for a total of 15,000 trees. All trees after the burn-in from each of the four independent runs were combined in PAUP* (Swofford 1993) to create one 50% majority consensus tree with support values indicative of posterior probability at each recovered node. For the ML and MLE analyses, the model of sequence evolution (MoSE) used was the one specified from the AIC indices from MODELTEST v3.7 (Posada, Crandall 1998; Posada, Crandall 2001; Posada, Buckley 2004), which was GTR+I+ Γ . Trees were rooted using *Entosiphon sulcatum* (GenBank Accession# FJ807254) since it was found to be the most basal euglenid in the results of Chapters 1 and 2. The kinetoplastid alignment was processed with ML using RAxML as described previously with the MoSE specified by Modeltest, GTR+ Γ , with 500 bootstrap replicates.

Analysis of positive selection

We used codeml in PAML v4.1 (Yang 1997; Yang 2007) to check for any instance of positive selection along any branches in the tree. This program allows comparisons of various models that observe the number of nonsynonymous (dN) mutations in the nucleotide level that would change the amino acid code of the protein sequence vs. the synonymous (dS) mutations that would not change the amino acid. At first the 60 euglenid EF1 α nucleotide sequences were processed using the null model (model=0) of evolution that specifies one dN/dS ratio for all branches in the tree, which output the estimated dN and dS for each branch given the data for the

fixed dN/dS ratio. For this data set codeml set the fixed dN/dS ratio to be 0.0256 and a total of 120 parameters were estimated using the null model. Then the same data set was processed through codeml using the alternate model hypothesis that dN/dS is different for every branch on the tree (model=1). A total of 237 parameters were estimated using this analysis. A likelihood ratio test (LRT; comparable to a chi squared (χ^2) test with the degrees of freedom equal to the differences in parameters estimated between the two analyses) was used to determine if there was a significant difference ($p > 0.05$) between the two models given the data set using the calculation: $\chi^2 = 2(\ln L_{\text{model1}} - \ln L_{\text{model0}})$ (Yang 1998; Yang 2001). The branches with any possible evidence of positive selection were determined and analyzed for strength of support of positive selection based on dN, dS and the dN/dS.

Results:

Multiple divergent copies of EF1 α in some euglenid species

There are multiple divergent versions of EF1 α in some euglenid lineages, but not in others (Fig. 4.3). *Cryptoglana* is nearly recovered as a monophyletic clade, although one extremely diverged gDNA sequence of *C. pigra* is found branching within a “crown group” of highly derived euglenid EF1 α sequences (Fig. 4.3). Within the main *Cryptoglana* cluster there are two cDNA sequences for both species of *Cryptoglana* (*C. pigra* and *C. skujai*), and instead of the copies from the same species branching more closely together, they branch as sister to a copy of EF1 α from the other species. The “crown group” that includes a version of *C. pigra* EF1 α seems to include a variety of taxa that are not recovered as a single clade in previous phylogenetic studies of the photosynthetic euglenids (Triemer et al. 2006; Kim, Shin 2008). Generally, the Eutreptiales are basal to the Euglenea, but they are instead present as polyphyletic in this “crown group” with gDNA EF1 α copies of *C. pigra* and *Euglena gracilis* as well as

cDNA from two polyphyletic *Monomorphina* species. These sequences appear to be highly divergent due to the long branch that leads from the other copies of *E. gracilis* EF1 α and several long branches within the group itself. Therefore, this *C. pigra* sequence is distinct and seemingly divergent from the aforementioned versions of *Cryptoglena* that are recovered together.

A phenomenon similar to the “recovered” *Cryptoglena* group is evident for the *Strombomonas* lineage, although the history of the gene duplication event does not appear to be as recent since other genera are included as sister to and between some of the *Strombomonas* groups of EF1 α recovered in various places in the tree (Fig. 4.3). However, it is interesting to note that the same clades are nearly always recovered where copies of *S. acuminata* strains are recovered as sister to a *S. napiformis* cDNA version of EF1 α in *Strombomonas* clade “A” and in *Strombomonas* clade “B” a separate copy of *S. acuminata* EF1 α is also recovered as sister to and different copy of *S. napiformis*. These two isoforms of the proteins are recovered as closely related, but separated by several *Trachelomonas* and *Euglena* species such that a clear duplication event between those species of *Strombomonas* cannot be assumed.

Although *S. ovalis* 1 in *Strombomonas* group “C” comes out as sister to *Monomorphina aenigmatica* 1, this close relationship is recovered consistently with both paralogs of EF1 α in both organisms. However, the *S. ovalis* 2 copy in the *Strombomonas* group “D” is sister to *S. borysteniensis* 3, which is sister to the *Strombomonas* group “C” *S. ovalis* 1 + *M. aenigmatica* 1 relationship. The one *Strombomonas* species that is not consistently recovered is *S. triquetra*. One version of EF1 α from *S. triquetra* branches as basal to one of the *S. acuminata* + *S. napiformis* clade in *Strombomonas* group “A”, whereas the other is basal to the *S. ovalis* + *M. aenigmatica* + *S. borysteniensis* clade in *Strombomonas* group “C”.

Another interesting line of evidence of gene duplications in the euglenid genera is that of the *Ploeotia* EF1 α sequences. The gDNA sequence of *P. oblonga* consistently branches as it is expected, basal to the euglenid tree within the Heteronematina (colorless euglenids) (Busse, Preisfeld 2003b; von der Heyden et al. 2004). However, the *P. oblonga* cDNA sequence is recovered as being nested within the Euglenea (phototrophic euglenids), which seems aberrant. It should be noted that the cDNA sequence of *P. oblonga* that was used in previous chapters to resolve deep level relationships of either EF1 α (Ch 1) or of the Euglenozoa (Ch 2) was always recovered as sister to the cDNA sequence of *P. costatum* sequence in a monophyletic clade. The latter was not included in this analysis because it was 99.7% identical in shared sites with *P. oblonga* 1 and therefore was culled from this analysis so as not to skew the tip-level relationships of the organisms by essentially having the same sequence twice in the same data set.

There were also several distinct sequences from *P. oblonga* gDNA acquired from cloning that were not included in this analysis because they were >95% identical to one another when their variable inserted sequences were removed (Table 4.1). The sequences clustered together within the basal Heteronematina during preliminary studies (data not shown) and this led to the removal of all but one to again avoid skewing the analyses by essentially including identical sequences. The inserted sequences were of variable length and the sequence within them was not conserved between different copies, but the positions of the insertions were conserved between these copies. Therefore, the inserted sequences are posited to be introns although this was not confirmed in this study.

Multiple similar copies of EF1 α in Trypanosomatid kinetoplastid species

Although the kinetoplastids also have evidence of gene duplication in their genomes, the gene duplications seem to be idiosyncratic and their derived lineages are not indicative of a

common duplication event with the divergent euglenid EF1 α copies in the shared ancestral euglenozoan ancestor. All kinetoplastid EF1 α sequences found in the four species with complete genomes sequences were found to be more closely related to other EF1 α sequences within the same kinetoplastid taxon (nearly 100% sequence identity of shared sites) than they were to possible orthologs or paralogs in the other kinetoplastid species (Table 4.2, Fig. 4.4). However, there does seem to be a similar history of the EF1 α sequences in the *Leishmania* sequences in which there are seven copies of EF1 α for both *L. major* and *L. infantum* and synteny is maintained among those seven copies, all of which reside on chromosome 17 for both species (Table 4.2). Synteny of varying copies of EF1 α is not maintained between the *Leishmania* and *Trypanosoma* lineages or even within these *Trypanosoma* lineages since *T. brucei* EF1 α sequences are mapped on chromosome 10, whereas *T. cruzi* maps to chromosome 5 and the *Leishmania* species' paralogs of EF1 α are mapped onto chromosome 17. Members of the genus *Trypanosoma* seem to have a dissimilar history either due to increased gene duplication and/or loss in one lineage or the other. Evidence for this comes from *T. brucei* which has three copies of EF1 α , whereas its sister taxon *T. cruzi* only has two copies (Table 4.2).

Inconclusive evidence of positive selection of EF1 α in Euglenida

Although there is some evidence of positive selection along branches in the euglenid EF1 α phylogeny, it is not strongly supported in the dN/dS ratios on the alternative model tree topology or statistically by the χ^2 test because it lacked power. When the topology was created with model 1 (dN/dS is different for every branch on the tree) there were many instances of possible evidence of positive selection. Unfortunately, these lacked support because in most cases there was either a zero value for dS for that given branch or both dN and dS values were zero. The strongest cases of possible positive selection is the branch leading from the

Heteronematina (basal colorless euglenids) to the group comprised of organisms within the Euglenea (phototrophic euglenids, with the exception of the cDNA copy of *Ploeotia oblonga*) with a very high dN (23.2) and a dS of 0 as well as the branch leading from *Euglena gracilis* genomic and cDNA sequences to the highly diverged “crown group” of euglenid EF1 α sequences, which includes the divergent gDNA copies of *Cryptoglana pigra* and *E. gracilis* as well as the cDNA sequences from *Monomorphina* sp., *Monomorphina pyrum* and the eutreptiales with a moderate amount of dN (9.5) and a dS of 0. Unfortunately, the LRT between the tree topologies of model 1 and model 0 was not significant ($p=0$), because although the χ^2 value was 424.157674, the degrees of freedom was 117 which minimized the power of the statistical test to distinguish and support significant differences between the two analyses.

Discussion:

Gene duplication of EF1 α seems to be only in the Euglenida lineage of the euglenozoans

There are multiple gene duplication events of EF1 α in the history of the Euglenida lineage. The most obvious evidence of a recent gene duplication event is in the ancestor of the *Cryptoglana* group. This is supported by the presence of two paralogs of mRNA (and one gDNA) in each of two species within that genus that grouped as closer to the ortholog of the EF1 α from the other species rather than to the paralog of EF1 α from the same taxon. This indicates that the ancestor *Cryptoglana* had both copies of EF1 α and that these two copies diverged from one another in the ancestor that gave rise to *C. pigra* and *C. skujai*. Both taxa received representative copies of both type 1 and type 2 EF1 α from their common ancestor. This would explain why the orthologs of type 1 ancestral *Cryptoglana* EF1 α are recovered as sister to each other as represented by the sequences from both species, as are the orthologs of type 2 ancestral *Cryptoglana* EF1 α .

A confounding point arises when we consider the highly divergent *C. pigra* gDNA EF1 α in the “crown group” euglenid EF1 α sequences. This genomic copy of *Cryptoglena* EF1 α sequence seems to be much more ancient as compared to the other *Cryptoglena* copies that cluster together. This is supported by the presence of EF1 α from the Eutreptiales as members of the “crown group”, which are thought to be basal to the Euglenia (Nudelman et al. 2006; Kim, Shin 2008). The long branch that extends from the *E. gracilis* clade to this “crown group” exhibits the possibility of positive selection based on its estimated dN of 9.5, suggesting that this “crown group” represents a version of euglenid EF1 α that resulted from a gene duplication event very early in the ancestor of the Euglenida. If selection pressure has been relaxed and it has since gathered enough non-synonymous mutations, it would now be extremely divergent from the other paralog(s) of euglenid EF1 α .

The versions of EF1 α in the *Strombomonas* lineages may represent at least one other gene duplication of EF1 α in the Euglenida. However, the presence of other organisms in and between these clusters of *Strombomonas* sequences supports the idea that this is not a duplication that is specific to the *Strombomonas* lineage. The *Monomorphina aenigmatica* sequence that arises with the *S. ovalis* and *S. borysteniensis* lends support that this gene duplication event happened in at least the most recent common ancestor (MRCA) of the *Monomorphina* and *Strombomonas* lineages, which seems to be fairly ancient relative to the phylogenies of extant lineages of the Euglenida. Also, because the *S. acuminata* and *S. borysteniensis* clades of the two types of EF1 α are separated by *Trachelomonas* and *Euglena* species, it seems that if this is a separate instance of gene duplication of EF1 α from the aforementioned *Strombomonas* gene duplication event, then this would have had to have occurred in the MRCA of those three genera. The fact that the *S. triquetra* divergent sequences of EF1 α cluster in both the former and the

latter clades of *Strombomonas* leads to the parsimonious explanation that there was only one gene duplication event involving the *Strombomonas*, *Monomorphina*, *Trachelomonas* and *Euglena* species involved and that the organisms that only have one copy have either lost the other or it has not been recovered with PCR and sequencing efforts. These options might also explain the other sequences whose single organism-level gene duplications cannot be resolved with the given data: *Euglena hiemalis*, *E. stellata* and *Lepocinclis tripteris*.

Finally, within the Euglenida it appears that there was either a single gene duplication event or an even more basal event occurred in the MRCA of the Euglenida. This conclusion is based on the EF1 α sequences recovered from the genus *Ploeotia*. The gDNA sequence of *Ploeotia* clusters within the Heteronematina, which have been supported to be basal to the Euglenida and are thought to have diverged prior to the Euglenea (Leander, Witek, Farmer 2001). Since the EF1 α cDNA versions of *P. oblonga* and *P. costatum* are recovered as nested within the Euglenea in the results from Chapters 1 and 2, we hypothesize that these divergent sequence types within the *Ploeotia* lineage are the result of an ancient gene duplication in the MRCA of the Euglenida or at least in the MRCA between the *Ploeotia* and the Euglenea lineages.

There are other possible explanations of these data; however, they are very unlikely. One alternate explanation is that there is a mechanism that edits the gDNA to create a seemingly phylogenetically divergent mRNA sequence. Although RNA editing is well documented in the mitochondria of the diplomonads and the kinetoplastids (Landweber, Gilbert 1994; Maslov et al. 1994; Maslov, Simpson 1994; Simpson et al. 2000; Marande, Lukes, Burger 2005; Roy et al. 2007), which are the other two main superclasses within the Euglenozoa with the euglenids, RNA editing of nuclear genes has not been previously reported in any organisms to date.

Therefore, it seemed highly unlikely that RNA editing was the explanation of the data found in this study. Also unlikely is the invoking rampant horizontal gene transfer (HGT) to explain why euglenid tip-level relationships and genera are not recovered according to previous studies (Leander, Farmer 2000; Leander, Witek, Farmer 2001; Brosnan et al. 2003; Brown, Zakrys, Farmer 2003; Brosnan et al. 2005; Milanowski et al. 2006; Triemer et al. 2006; Ciugulea et al. 2007; Ciugulea et al. 2008; Kim, Shin 2008; Kosmala et al. 2009). In order for this option to be feasible the HGT donor(s) would have to be euglenids as well, and there is no known direct mechanism or documentation of nuclear gene replacement in such closely related organisms.

The gene duplication events of the divergent EF1 α -types appear to only be represented in the Euglenida lineage and not in the Kinetoplastida and therefore, not in the MRCA of the Euglenozoa since those are the only two superclasses in the Euglenozoa that retain endogenous EF1 α (Ch 2). The high level of identity between the various copies of EF1 α in the trypanosomatids such that no distinct and divergent clone types are present within any given species is a very discernible difference between the Euglenida and Kinetoplastida histories of EF1 α . This could be seen as evidence of concerted evolution of the EF1 α in the kinetoplastid lineages, but that seems an unlikely explanation based on the presence of nearly identical multiple copies of EF1 α in lineages that are very distantly related to the Euglenozoa (Kidou, Ejiri 1998; Vinkenoog et al. 1998). The nearly identical sequences of multiple copies of EF1 α in the genomes of different organisms is what allows it to be functionally considered an effectively single copy nuclear-encoded gene, one reason why it may prove to be useful at recovering deep-level phylogenies. More likely is the alternative that the copies present in the Kinetoplastida do not result from the gene duplication event that led to the highly divergent sequence types of EF1 α evident in the Euglenida. This would suggest that the timing of the gene duplication even

that led to those divergent EF1 α sequence types was in the MRCA of the Euglenida or at least very early in the evolution of the euglenid lineage.

The divergent euglenid EF1 α gene duplicates are most likely still functional

There were only a few instances where there might have been a difference in dN/dS that hinted at positive selection in the divergent EF1 α gene sequences within the Euglenida; however, because there were only zero values in the dS category, the results were still inconclusive. EF1 α is a very conserved gene based on its important function in eukaryotic translation and because all copies are still relatively similar in their sequences without any nonsense mutations, this supported the fact that the divergent copies are still expressed and functional within the organisms. This is definitely supported in the cases where the sequences were derived from total RNA, and the differential recovery of clones associated with different copy types could be indicative of varying transcript levels between the copies within the organism (Table 4.1). There is also no evidence of the gDNA copies being pseudogenes because no stop codons were found that would have caused early termination of the coding of the gene.

Conclusion:

EF1 α is still useful in recovering phylogenies of organisms at certain levels, but extreme care must be taken to be sure that orthologs and paralogs are recognized and correctly used in the reconstruction such that the gene family history does not obscure the species relationships, such as in the Euglenida. The fact that EF1 α is so conserved across the eukaryotic tree of life suggests that it can be used to resolve deep-level clades whether paralogs or orthologs are used if a study focuses on deeper-level questions and not more recent phylogenetic relationships. Although the “gene tree” of EF1 α did not match the previously supported “species tree” of euglenids, the gene can still be used to recover specific tip-level relationships of the euglenids

because there are several independent gene duplication events that are specific to certain lineages such as *Cryptoglena* and possibly *Strombomonas*. However, this is unclear because other genera may also have the orthologs, but they simply were not recovered in the PCR and cloning process. This would have to be examined further with extensive amounts genome sequencing or Southern blots and biochemical analyses across the breadth of the Euglenida.

References:

- Akamine, RN, CE Winter. 2008. *Oscheius tipulae* as an example of eEF1A gene diversity in nematodes. *Journal of Molecular Evolution* 67:278-290.
- Baldauf, SL. 2003. The deep roots of eukaryotes. *Science* 300:1703-1706.
- Baldauf, SL, WF Doolittle. 1997. Origin and evolution of the slime molds (Mycetozoa). *Proceedings of the National Academy of Sciences of the United States of America* 94:12007-12012.
- Baldauf, SL, JD Palmer. 1993. Animals and Fungi Are Each Others Closest Relatives - Congruent Evidence from Multiple Proteins. *Proceedings of the National Academy of Sciences of the United States of America* 90:11558-11562.
- Baldauf, SL, AJ Roger, I Wenk-Siefert, WF Doolittle. 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* 290:972-977.
- Breglia, SA, CH Slamovits, BS Leander. 2007. Phylogeny of phagotrophic euglenids (Euglenozoa) as inferred from HSP90 gene sequences. *Journal of Eukaryotic Microbiology* 54:86-92.
- Brosnan, S, PJP Brown, MA Farmer, RE Triemer. 2005. Morphological separation of the euglenoid genera *Trachelomonas* and *Strombomonas* (Euglenophyta) based on lorica development and posterior strip reduction. *Journal of Phycology* 41:590-605.
- Brosnan, S, W Shin, KM Kjer, RE Triemer. 2003. Phylogeny of the photosynthetic euglenophytes inferred from the nuclear SSU and partial LSU rDNA. *International Journal of Systematic and Evolutionary Microbiology* 53:1175-1186.
- Brown, PJP, B Zakrys, MA Farmer. 2003. Plastid morphology, ultrastructure, and development in *Colacium* and the loricate euglenophytes (Euglenophyceae). *Journal of Phycology* 39:115-121.
- Burki, F, J Pawlowski. 2006. Monophyly of Rhizaria and multigene phylogeny of unicellular bikonts. *Molecular Biology and Evolution* 23:1922-1930.
- Burki, F, K Shalchian-Tabrizi, J Pawlowski. 2008. Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. *Biology Letters* 4:366-369.
- Busse, I, A Preisfeld. 2002. Phylogenetic position of *Rhynchopus* sp. and *Diplonema ambulator* as indicated by analyses of euglenozoan small subunit ribosomal DNA. *Gene* 284:83-91.
- Busse, I, A Preisfeld. 2003a. Application of spectral analysis to examine phylogenetic signal among euglenid SSU rDNA data sets (Euglenozoa). *Organisms Diversity & Evolution* 3:1-12.
- Busse, I, A Preisfeld. 2003b. Discovery of a group I intron in the SSU rDNA of *Ploeotia costata* (Euglenozoa). *Protist* 154:57-69.

- Ciugulea, I, MA Nudelman, S Brosnan, RE Triemer. 2007. Phylogeny of the euglenoid loricate genera *Trachelomonas* and *Strombomonas* (Euglenophyta) inferred from nuclear ssu and lsu rDNA. *Journal of Phycology* 43:196.
- Ciugulea, I, MA Nudelman, S Brosnan, RE Triemer. 2008. Phylogeny of the euglenoid loricate genera *Trachelomonas* and *Strombomonas* (Euglenophyta) inferred from nuclear SSU and LSU rDNA. *Journal of Phycology* 44:406-418.
- de Queiroz, A, J Gatesy. 2007. The supermatrix approach to systematics. *Trends in Ecology & Evolution* 22:34-41.
- Drummond, AJ, B Ashton, M Cheung, J Heled, M Kearse, R Moir, S Stones-Havas, T Thierer, A Wilson. 2009. Geneious: Biomatters Ltd.
- Edgar, RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792-1797.
- Fiore-Donno, AM, CD Berney, J Pawlowski, SL Baldauf. 2005. Higher-order phylogeny of plasmodial slime molds (Myxogastria) based on elongation factor 1-A and small subunit rRNA gene sequences. *Journal of Eukaryotic Microbiology* 52:201-210.
- Gile, GH, PJ Keeling. 2008. Nucleus-encoded periplastid-targeted EFL in chlorarachniophytes. *Molecular Biology and Evolution* 25:1967-1977.
- Gile, GH, NJ Patron, PJ Keeling. 2006. EFL GTPase in cryptomonads and the distribution of EFL and EF-1 alpha in chromalveolates. *Protist* 157:435-444.
- Huelsenbeck, JP, F Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.
- Keeling, PJ, Y Inagaki. 2004. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1 alpha. *Proceedings of the National Academy of Sciences of the United States of America* 101:15380-15385.
- Kidou, S, S Ejiri. 1998. Isolation, characterization and mRNA expression of four cDNAs encoding translation elongation factor 1A from rice (*Oryza sativa* L.). *Plant Molecular Biology* 36:137-148.
- Kim, JI, W Shin. 2008. Phylogeny of the euglenales inferred from plastid LSU rDNA sequences. *Journal of Phycology* 44:994-1000.
- Kosmala, S, A Karnkowska-Ishikawa, R Milanowski, J Kwiatowski, B Zakrys. 2009. Phylogeny and systematics of *Euglena* (Euglenaceae) species with axial, stellate chloroplasts based on morphological and molecular data-new taxa, emended diagnoses and epitypifications. *Journal of Phycology* 45:464-481.

- Landweber, LF, W Gilbert. 1994. Phylogenetic Analysis of Rna Editing - a Primitive Genetic Phenomenon. *Proceedings of the National Academy of Sciences of the United States of America* 91:918-921.
- Leander, BS, MA Farmer. 2000. Comparative morphology of the euglenid pellicle. I. Patterns of strips and pores. *Journal of Eukaryotic Microbiology* 47:469-479.
- Leander, BS, MA Farmer. 2001. Evolution of Phacus (Euglenophyceae) as inferred from pellicle morphology and SSU rDNA. *Journal of Phycology* 37:143-159.
- Leander, BS, RP Witek, MA Farmer. 2001. Trends in the evolution of the euglenid pellicle. *Evolution* 55:2215-2235.
- Leedale, G. 1967. Euglenoid Flagellates. 1-73.
- Li, L, CJ Stoeckert, DS Roos. 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research* 13:2178-2189.
- Linton, EW, MA Nudelman, V Conforti, RE Triemer. 2000. A molecular analysis of the euglenophytes using SSU rDNA. *Journal of Phycology* 36:740-746.
- Marande, W, J Lukes, G Burger. 2005. Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. *Eukaryotic Cell* 4:1137-1146.
- Maslov, DA, HA Avila, JA Lake, L Simpson. 1994. Evolution of Rna Editing in Kinetoplastid Protozoa. *Nature* 368:345-348.
- Maslov, DA, L Simpson. 1994. Rna Editing and Mitochondrial Genomic Organization in the Cryptobiid Kinetoplastid Protozoan Trypanoplasma-Borrelia. *Molecular and Cellular Biology* 14:8174-8182.
- Milanowski, R, S Kosmala, B Zakrys, J Kwiatowski. 2006. Phylogeny of photosynthetic euglenophytes based on combined chloroplast and cytoplasmic SSU rDNA sequence analysis. *Journal of Phycology* 42:721-730.
- Miller, MA, MT Holder, R Vos, PE Midford, T Liebowitz, L Chan, P Hoover, T Warnow. 2009. The CIPRES Portals.
- Mochida, K, T Matsubara. 2002. Molecular cloning of an elongation factor 1 alpha and its mRNA localization in testis of the Nile tilapia *Oreochromis niloticus*. *Fisheries Science* 68:830-837.
- Montandon, PE, E Stutz. 1990. Structure and Expression of the *Euglena gracilis* Nuclear Gene Coding for the Translation Elongation-Factor Ef-1-Alpha. *Nucleic Acids Research* 18:75-82.

- Nudelman, MA, PI Leonardi, V Conforti, MA Farmer, RE Triemer. 2006. Fine structure and taxonomy of *Monomorpha aenigmatica* comb. nov (Euglenophyta). *Journal of Phycology* 42:194-202.
- Posada, D, TR Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53:793-808.
- Posada, D, KA Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817-818.
- Posada, D, KA Crandall. 2001. Selecting the best-fit model of nucleotide substitution. *Systematic Biology* 50:580-601.
- Ronquist, F, JP Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
- Roy, J, D Faktorova, J Lukes, G Burger. 2007. Unusual mitochondrial genome structures throughout the Euglenozoa. *Protist* 158:385-396.
- Simpson, AGB, AJ Roger. 2004. Protein phylogenies robustly resolve the deep-level relationships within Euglenozoa. *Molecular Phylogenetics and Evolution* 30:201-212.
- Simpson, L, OH Thiemann, NJ Savill, JD Alfonzo, DA Maslov. 2000. Evolution of RNA editing in trypanosome mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 97:6986-6993.
- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.
- Stamatakis, A, P Hoover, J Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Systematic Biology* 57:758-771.
- Swofford, DL. 1993. PAUP - A computer program for phylogenetic inference using Maximum Parsimony. *Journal of General Physiology* 102:A9-A9.
- Tanabe, Y, M Saikawa, MM Watanabe, J Sugiyama. 2004. Molecular phylogeny of Zygomycota based on EF-1 alpha and RPB1 sequences: limitations and utility of alternative markers to rDNA. *Molecular Phylogenetics and Evolution* 30:438-449.
- Triemer, RE, MA Farmer. 2007. A decade of euglenoid molecular phylogenetics. In: J Brodie, J Lewis, editors. *Unraveling the algae: the past, present and future of algal systematics: Systematics Association Series*, CRC Press. p. 315-330.
- Triemer, RE, E Linton, W Shin, A Nudelman, A Monfils, M Bennett, S Brosnan. 2006. Phylogeny of the euglenales based upon combined SSU and LSU rDNA sequence comparisons and description of *Discoplastis* gen. nov (Euglenophyta). *Journal of Phycology* 42:731-740.

- Vinkenoog, R, MA Speranca, O van Breemen, J Ramesar, DH Williamson, PB Ross-MacDonald, AW Thomas, CJ Janse, HA del Portillo, AP Waters. 1998. Malaria parasites contain two identical copies of an elongation factor 1 alpha gene. *Molecular and Biochemical Parasitology* 94:1-12.
- von der Heyden, S, EE Chao, K Vickerman, T Cavalier-Smith. 2004. Ribosomal RNA phylogeny of bodonid and diplomonid flagellates and the evolution of euglenozoa. *Journal of Eukaryotic Microbiology* 51:402-416.
- Xavier, MB, CSR Mainardespinto, M Takino. 1991. *Euglena sanguinea* Ehrenberg bloom in a fish-breeding tank (Pindamonhangaba, Sao-Paulo, Brazil). *Archiv Fur Hydrobiologie*:133-142.
- Yang, ZH. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences* 13:555-556.
- Yang, ZH. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular Biology and Evolution* 15:568-573.
- Yang, ZH. 2001. Adaptive molecular evolution. In: D Balding, M Bishop, C Cannings, editors. *Handbook of Statistical Genetics*. New York: Wiley. p. 327-350.
- Yang, ZH. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24:1586-1591.
- Zakrys, B. 1988. The nuclear DNA level as a potential taxonomic character in *Euglena* Ehr. (Euglenophyceae). *Arch. Hydrobiol. Suppl.* 78:483-504.

Table 4.1. Euglenid organism information for EF1 α sequences used in this study. Bold in accession # column denotes new sequences obtained in this study. Only sequencing information for organisms we performed are included in this table. Copy # was arbitrarily to differentiate between and denote the number of distinct sequence types (<95% sequence identity of shared sites) for a given strain of species. The type of total nucleic acid that was originally extracted from the organisms is noted under NA Source; messenger RNA (mRNA) or genomic DNA (gDNA). The method that led directly to sequencing (Seq. Method) was either PCR purification (PCR), cloning, or 5' rapid amplification of cDNA ends (RACE). The number of clones (#clones) that were sequenced and used to create the final contiguous sequence (contig) are listed. The asterisk next to the single *Ploeotia oblonga* gDNA clone represents that there were 3 clones from the same cloning reaction that had conserved inserted sequence (assumed, but not confirmed to be introns) placement, but insertion sizes were variable; however, they were >95% identical when inserted sequences were removed. Those three clones and a separate *P. oblonga* gDNA clone that was <95% identical with the other three (when inserted sequences were excluded) formed a clade which was recovered basal to the Euglenea, as was expected of representative sequences of the Heteronematina based on prior studies.

"Group(s)"	Species	copy#	NA Source	Seq. Method	#clones	Strain	Accession #
Euglenea/Colacium	<i>Colacium elongatum</i>	1	mRNA	cloning	10	M11	apply for
Euglenea/Colacium	<i>Colacium</i> sp.	1	mRNA	PCR+cloning	3	NJ	apply for
Euglenea/Colacium	<i>Colacium vesiculosum</i>	1	mRNA	PCR	n/a	UW Lazienki	apply for
Euglenea/Cryptoglena	<i>Cryptoglena pigra</i>	1	mRNA	cloning	5	CCAP 1212/1	apply for
		2	mRNA	cloning	4		apply for
		3	gDNA	cloning	2		apply for
		4	gDNA	cloning	1		apply for
Euglenea/Cryptoglena	<i>Cryptoglena skujai</i>	1	mRNA	cloning	4	SAG 10.88	apply for
		2	mRNA	cloning	5		apply for
Euglenea/Discoplastis	<i>Discoplastis adunca</i>	1	mRNA	cloning	6	ASW 08039	apply for
Euglenea/Euglena	<i>Euglena clara</i>	1	mRNA	cloning	3	SAG 25.98	apply for
Euglenea/Euglena	<i>Euglena deses</i>	1	mRNA	cloning	6	SAG 1224-23	apply for
Euglenea/Euglena	<i>Euglena exilis</i>	1	mRNA	cloning	4	SAG 1224-11a	apply for
Euglenea/Euglena	<i>Euglena geniculata</i>	1	mRNA	cloning	10	SAG 1224-21	apply for
Euglenea/Euglena	<i>Euglena gracilis</i>	1	mRNA	cloning	4	Z strain	apply for
		2	gDNA	cloning	5		apply for
		3	gDNA	cloning	3		apply for
Euglenea/Euglena	<i>Euglena granulata</i>	1	mRNA	cloning	8	OR5	apply for
Euglenea/Euglena	<i>Euglena hiemalis</i>	1	mRNA	cloning	1	CCAP 1224/35	apply for
		2	mRNA	cloning	3		apply for
Euglenea/Euglena	<i>Euglena stellata</i>	1	mRNA	cloning	2	IAM E-11	apply for
		2	mRNA	cloning	1		apply for
Euglenea/Euglena	<i>Euglena viridis</i>	1	mRNA	PCR	n/a	NJ001	apply for
Euglenea/Eutreptiales	<i>Eutreptia pertyi</i>	1	mRNA	cloning	4	UTEX LB 1290	apply for

Euglenea/Eutreptiales	<i>Eutreptiella pomquetensis</i>	1	mRNA	cloning	1	CCMP 1491	apply for
Euglenea/ <i>Lepocinclis</i>	<i>Lepocinclis acus</i> var. <i>major</i>	1	mRNA	PCR+cloning	1	UTEX LB 1316	apply for
Euglenea/ <i>Lepocinclis</i>	<i>Lepocinclis ovum</i>	1	mRNA	cloning	4	SAG 1244-8	apply for
Euglenea/ <i>Lepocinclis</i>	<i>Lepocinclis steinii</i>	1	mRNA	cloning	3	UTEX 523	apply for
Euglenea/ <i>Lepocinclis</i>	<i>Lepocinclis tripteris</i>	1	mRNA	cloning	10	UTEX 1311	apply for
		2	mRNA	cloning	3		apply for
Euglenea/ <i>Monomorphina</i>	<i>Monomorphina aenigmatica</i>	1	mRNA	cloning	4	UTEX 1284	apply for
		2	mRNA	cloning	1		apply for
Euglenea/ <i>Monomorphina</i>	<i>Monomorphina pyrum</i>	1	mRNA	cloning	6	UTEX 2354	apply for
Euglenea/ <i>Monomorphina</i>	<i>Monomorphina</i> sp.	1	mRNA	cloning	2	MI 73	apply for
Euglenea/ <i>Phacus</i>	<i>Phacus orbicularis</i>	1	mRNA	PCR	n/a	ACOI 996	apply for
Euglenea/ <i>Phacus</i>	<i>Phacus pusillus</i>	1	mRNA	cloning	3	UTEX LB 1282	apply for
Euglenea/ "loricate" clade/ <i>Strombomonas</i>	<i>Strombomonas acuminata</i>	1	mRNA	RACE	n/a	ACOI2467	apply for
Euglenea/ "loricate" clade/ <i>Strombomonas</i>	<i>Strombomonas acuminata</i>	1	mRNA	cloning	3	SAG 1280-1	apply for
		2	mRNA	cloning	4		apply for
Euglenea/ "loricate" clade/ <i>Strombomonas</i>	<i>Strombomonas borysteniensis</i>	1	mRNA	cloning	1	NJ Stromb 10	apply for
		2	gDNA	cloning	1		apply for
		3	mRNA	cloning	1		apply for
Euglenea/ "loricate" clade/ <i>Strombomonas</i>	<i>Strombomonas napiformis</i>	1	mRNA	cloning	2	NJ Strombo 2	apply for
		2	mRNA	cloning	1		apply for
		3	gDNA	cloning	3		apply for
Euglenea/ "loricate" clade/ <i>Strombomonas</i>	<i>Strombomonas ovalis</i>	1	mRNA	cloning	6	NJ S115	apply for
		2	mRNA	cloning	3		apply for
Euglenea/ "loricate" clade/ <i>Strombomonas</i>	<i>Strombomonas triquetra</i>	1	mRNA	cloning	4	NJ S116	apply for
		2	mRNA	cloning	3		apply for
Euglenea/ "loricate" clade/ <i>Trachelomonas</i>	<i>Trachelomonas conradi</i>	1	mRNA	PCR+cloning	1	UTEX 1313	apply for
Euglenea/ "loricate" clade/ <i>Trachelomonas</i>	<i>Trachelomonas echinata</i>	1	mRNA	PCR+cloning	2	SAG 1283-22	apply for
Euglenea/ "loricate" clade/ <i>Trachelomonas</i>	<i>Trachelomonas lefevrei</i>	1	mRNA	PCR+ cloning	1	SAG 1283-10	apply for
Euglenea/ "loricate" clade/ <i>Trachelomonas</i>	<i>Trachelomonas planctonica</i>	1	mRNA	cloning	9	Woongghi Trach 1(WT1)	apply for
Euglenea/ "loricate" clade/ <i>Trachelomonas</i>	<i>Trachelomonas similis</i>	1	mRNA	cloning	2	SAG 1283-14	apply for
Euglenea/ "loricate" clade/ <i>Trachelomonas</i>	<i>Trachelomonas volvocina</i>	1	mRNA	PCR+cloning	2	UTEX 1327	apply for
Heteronematina/ <i>Entosiphon</i>	<i>Entosiphon</i> sp.	1	mRNA	cloning	1	Carolina Biological	apply for
Heteronematina/ <i>Entosiphon</i>	<i>Entosiphon sulcatum</i>	1	gDNA	-	-		FJ807254
Heteronematina/ <i>Peranema</i>	<i>Peranema trichophorum</i>	1	mRNA	cloning	1	CCAP 1260/1b	apply for
		2	mRNA	-	-		FJ807243
Heteronematina/ <i>Ploeotia</i>	<i>Ploeotia oblonga</i>	1	mRNA	cloning	4	ATCC 50573	apply for
	<i>Ploeotia oblonga</i>	2	gDNA	cloning	1*	ATCC 50573	apply for

Table 4.2. Kinetoplastid EF1 α sequence recovery from OrthoMCL utilized in the TriTrypDB v1.1. Copy#=number of EF1 α copies detected by OrthoMCL in the haploid genomes of each kinetoplastid species; chromosome=the chromosome where all copies of EF1 α were found in each kinetoplastid species; % identical sites=number of identical sites shared across the entire gapped alignment of 1352bp for all copies in each kinetoplastid species; sequence comments=any distinguishable differences between copies; Accession #=TriTrypDB identification information.

Species	copy#	chromosome	% identical sites	sequence comments	Accession #
<i>Leishmania infantum</i>	7	17	98.6%	n/a	LinJ17_V3.0090
					LinJ17_V3.0100
					LinJ17_V3.0110
					LinJ17_V3.0170
					LinJ17_V3.0180
					LinJ17_V3.0190
					LinJ17_V3.0200
<i>Leishmania major</i>	7	17	99.8%	n/a	LmjF17.0080
					LmjF17.0081
					LmjF17.0082
					LmjF17.0083
					LmjF17.0084
					LmjF17.0085
					LmjF17.0086
<i>Trypanosoma brucei</i> (TREU927 strain)	3	10	100.0%		Tb10.70.5650
					Tb10.70.5670
				missing 1-303bp	Tb10.70.5680
<i>Trypanosoma cruzi</i> (CL Brener Esmeraldo strain)	2	5	98.3%		Tc00.1047053511369.10
				missing 1-21bp	Tc00.1047053511369.20

Figure 4.1. The euglenid EF1 α interleaved alignment. The basal colorless (Heteronematina) are listed first and then euglenids based on their sequence identity and not previously recovered species-level relationships. The column positions, from 1-1167, are given in ten base increments at the top of each page. The number of residues from the original sequence is given in brackets to the right of each line corresponding to its respective OTU.

	120	130	140	150	160	170	180	190	200	210	220	
Entosiphon_sulcatum_FJ807254	CATCGAGAAG	TTTCGAGAAGG	AGGCCGCTGA	GATGGGCAAG	GCCAGCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGACA	[168]
Entosiphon_sp	CATCGAGAAG	TTTCGAGAAGG	AGGCCGCTGA	GATGGGCAAG	GCCAGCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGACA	[220]
Ploeotia_oblonga_gDNA_2	CATCGAGAAG	TTTCGAGAAGG	AGGCCGCTGA	GATGGGCAAG	GCTTCATTCA	AGTACGCTTG	GGTGTTAGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGACA	[220]
Peranema_trichophorum_FJ807243_2	CaTcGaaAaAg	TTCgAGaAaGg	AaGcAGcTcGa	gaTgggAaAa	GcCtCaTtCa	AaTaCgCtTg	GgTcTtGgAc	AaAcTYAaGg	CcGaAaGaGa	AaGaGgTaTc	AcCaTcGaCa	[214]
Peranema_trichophorum_2	CATCGAAAAAG	TTTCGAGAAGG	AAGCAGCTGA	GATGGGTAAG	GATCGCTTCA	AATACGCTTG	GGTCTTGGAC	AAACTCAAGG	CAGAAGCTGA	AAGAAGGTATC	ACCATCGACA	[220]
Euglena_stellata_2	-----	TTTCGAGAAGG	ARGCGGSCGA	GATSGGTAAA	GGTTCCTTCA	AGTATGCCTG	GGTCTTGGAT	AAGCTGAAGG	CCGAGCGTGA	ACGTGGTATC	ACCATCGATA	[100]
Euglena_hiemalis_1	CATTGAGAAG	TTTCGAGAAGG	AGGCCAGCTGA	GATGGGCAAG	GGTTCCTTCA	AGTATGCCTG	GGTCTTGGAT	AAGCTTAAGG	CCGAGCGTGA	ACGTGGTATC	ACCATCGACA	[220]
Euglena_hiemalis_2	CACCGAGAAA	TTTCGAGAAGG	AGGCCGCTGA	GATGGGCAAG	GGCTCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTTAAGG	CCGAGCGTGA	ACGTGGTATC	ACCATCGATA	[220]
Euglena_exilis	-----	TTTCGAGAAGG	AGGCCGCTGA	GATGGGTAAG	GGATCCTTCA	AGTACGCTTG	GGTCTTGGAC	AAGTTAAAGG	CTGAGCGTGA	ACGTGGTATC	ACCATCGATA	[100]
Strombomonas_triquetra_1	AATTGAGAAA	TTTGAGAAGG	AAGCCTCTGA	AATGGGCAAG	GGGTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAACTTAAGG	CTGAGCGTGA	ACGTGGTATC	ACCATCGATA	[220]
Strombomonas_napiformis_1	AATTGAGAAG	TTTGAGAAGG	AGGCATCTGA	GATGGGCAAG	GGGTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTCAAGG	CTGAGCGTGA	ACGTGGTATC	ACCATCGACA	[220]
Strombomonas_acuminata_ACOI_2467	AATTGAGAAA	TTTGAGAAGG	AAGCATCTGA	AATGGGCAAG	GGGTCCTTCA	AATATGCTTG	GGTCTTGGAC	AAGCTTAAGG	CTGAGCGTGA	ACGTGGTATC	ACCATCGACA	[220]
Strombomonas_acuminata_SAG1280_1_1	-----GG	TTTCGAGAAGG	AGGCCGATGA	GATGGGCAAG	GGGTCCTTCA	AATATGCTTG	GGTCTTGGAC	AAGCTTAAGG	CTGAGCGTGA	ACGTGGTATC	ACCATCGACA	[102]
Phacus_orbicularis	CATCGAAAAAG	TTTGAGAAGG	AAGCCTCTGA	CATGGGTAAG	GGTTCCTTCA	AGTACGCTTG	GGTGTGGAC	AAGCTCAAGG	CCGAACGTGA	GCAGCGGTATC	ACCATCGACA	[178]
Euglena_stellata_2_2	-----	--GAGAAGG	AGGCCGGTGA	GATGGGTAAG	GGTTCCTTCA	AGTACGCTTG	GGTGTAGAC	AAGCTGAAGG	CTGAGCGTGA	ACGTGGTATC	ACCATCGATA	[97]
Trachelomonas_similis	-----	TTTCGAGAAGG	AGGCCGCTGA	GATGGGCAAG	GGTTCCTTCA	AATATGCCTG	GGTGCTGGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[100]
Trachelomonas_volvocina	CATTGAGAAA	TTTCGAGAAGG	AAGCCTCTGA	AATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGTGGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Euglena_clara	-----	TTTCGAGAAGG	AGGCCGCTGA	GATGGGCAAG	GGGTCCTTCA	AGTACGCTTG	GGTGCTGGAT	AAGCTGAAGG	CAGAGCGTGA	GCAGCGGTATC	ACCATCGACA	[100]
Euglena_deses	CATCGAAAAAG	TTTCGAGAAGG	AGGCCCTCTGA	GATGGGCAAG	GGTTCCTTCA	AATACGCTTG	GGTGCTGGAT	AAGCTMAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGACA	[220]
Lepocinclis_ovum	-----	TTTCGAGAAGG	ARGCGGCTGA	GATSGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGTGGAC	AAGCTCAAGG	CTGAACGTGA	GCAGCGGTATC	ACCATCGATA	[100]
Euglena_granulata	CATCGAGAAG	TTTCGAGAAGG	AGGCTTCTGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTCTTGGAC	AAGCTTAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Trachelomonas_planctonica	-----GG	TTTCGAGAAGG	AGGCCGGTGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGTGGAC	AAGCTCAAGG	CAGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[102]
Trachelomonas_lefevrei	GATCGAGAAG	TTTCGAGAAGG	AGGCCGCCGA	GATGGGCAAG	GGGTCCTTCA	AGTACGCTTG	GGTGTGGAC	AAGCTCAAGG	CAGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Trachelomonas_echinata	CATCGAAAAAG	TTTCGAAAAAGG	AGGCCACCGGA	TATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTCTGGAC	AAGCTTAAGG	CAGAGCGTGA	ACGTGGTATC	ACCATCGATA	[220]
Discoplastis_adunca	CATCGAGAAG	TTTGAGAAGG	AGGCCCTCTGA	AATGGGCAAG	AGCTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Lepocinclis_trieteris_1	CATTGAGAAG	TTTGAGAAGG	AAGCCAATGA	CATGGGTAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Lepocinclis_acus_var_major	CATTGAGAAG	TTTGAGAAGG	AGGCCAATGA	CATGGGTAAG	GGTTCCTTCA	AGTACGCTTG	GGTGTGGAC	AAGCTCAAGG	CAGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Lepocinclis_steinii	-----	TTTCGAGAAGG	AGGCCGGTGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTCAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[100]
Monomorpha_aenigmatica_2	AATCGAAAAAG	TTTCGAGAAGG	AGGCCACCGGA	CATGGGTAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Strombomonas_ovalis_2	-----GG	TTTCGAGAAGG	AGGCCGGTGA	GATGGGTAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[102]
Strombomonas_borysteniensis_3	-----GG	AGGCCGGTGA	GATGGGTAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[92]	
Phacus_pusillus	CATTGAGAAG	TTTCGAGAAGG	AGGCCAATGA	CATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Lepocinclis_trieteris_2	-----GG	TTTCGAGAAGG	AGGCCGGTGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[102]
Trachelomonas_conradi	CATCGAGAAG	TTTCGAGAAGG	AAGCCTCTGA	GATGGGTAAG	GGTTCCTTCA	AATATGCCTG	GGTGCTGGAC	AAGCTCAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Monomorpha_sp	CATCGAGAAG	TTTCGAGAAGG	AGGCCACTGA	GATGGGTAAG	GGGTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Euglena_gracilis_gDNA_3	CATTGAGAAG	TTTCGAGAAGG	AGGCCACCGGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Eutreptiella_pomquetensis	CATTGAGAAG	TTTCGAGAAGG	AGGCCACCGGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Eutreptia_pertyi	CATTGAGAAG	TTTCGAGAAGG	AGGCCACCGGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Monomorpha_pyrum	CATTGAGAAG	TTTCGAGAAGG	AGGCCACCGGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Cryptoglena_pigra_gDNA_4	CATCGAGAAG	TTTCGAGAAGG	AAGCCGCCGA	GATGGGTAAG	GGTTCCTTCA	AGTACGCCG	GGTGCTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Euglena_gracilis_gDNA_2	CATCGAGAAG	TTTCGAGAAGG	AAGCCGCCGA	GATGGGTAAG	GGTTCCTTCA	AGTACGCCG	GGTGCTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Euglena_gracilis_1	CATCGAGAAG	TTTCGAGAAGG	AAGCCGCCGA	GATGGGTAAG	GGTTCCTTCA	AGTACGCCG	GGTGCTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Ploeotia_oblonga_1	CATCGAGAAG	TTTCGAGAAGG	AAGCCACCGGA	TATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Euglena_geniculata	CATCGAGAAG	TTTCGAGAAGG	AAGCCACCGGA	TATGGGCAAG	GGTTCCTTCA	AGTACGCTTG	GGTGCTGGAC	AAGCTGAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Cryptoglena_pigra_2	GATCGAGAAA	TTTCGAGAAGG	AAGCCTCTGA	GATGGGCAAG	GGCTCTTCA	AGTACGCCG	GGTGCTGGAC	AAGCTCAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Cryptoglena_pigra_gDNA_3	GATCGAGAAA	TTTCGAGAAGG	AAGCCTCTGA	GATGGGTAAG	GGCTCGTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTCAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGACA	[220]
Cryptoglena_skujai_2	GATCGAGAAA	TTTCGAGAAGG	AAGCCTCTGA	GATGGGCAAG	GGCTCGTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTCAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGACA	[220]
Cryptoglena_pigra_1	CATCGAGAAG	TTTGAGAAGG	AAGCGTCGGA	AATGGGCAAG	GGTTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAACTCAAGG	CTGAGCGTGA	ACGTGGTATC	ACCATCGATA	[220]
Cryptoglena_skujai_1	CATCGAGAAG	TTTCGAGAAGG	AAGCATCTGA	GATGGGCAAG	GGTTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTCAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Colacium_vesiculosum	CATCGAGAAG	TTTCGAGAAGG	AGGCCAAGTGA	AATGGGCAAG	GGCTCTTCA	AGTATGCCTG	GGTGTGGAC	AAGCTCAAGG	CTGAGCGTGA	ACGTGGTATC	ACCATCGACA	[179]
Colacium_sp	CATTGAGAAG	TTTCGAGAAGG	AGGCCAAGTGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCCG	GGTGTGGAC	AAGCTCAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGACA	[220]
Colacium_elongatum	TATCGAGAAG	TTTCGAGAAGG	AGGCCAAGTGA	GATGGGCAAG	GGTTCCTTCA	AGTACGCCG	GGTGTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Strombomonas_triquetra_2	CATCGAAAAAG	TTTCGAAAAAGG	AGGCCCTCTGA	GATGGGCAAG	GGGTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Strombomonas_ovalis_1	TATCGAAAAAG	TTTCGAAAAAGG	AAGCCTCTGA	TATGGGCAAG	GGTTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTTAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Monomorpha_aenigmatica_1	-----GG	AGGCCGCCGA	GATGGGCAAG	GGTTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTTAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[92]	
Strombomonas_borysteniensis_1	-----	TTTCGAGAAGG	AGGCCGCCGA	GATGGGCAAG	GGTTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[100]
Strombomonas_borysteniensis_gDNA_2	TATCGAAAAAG	TTTCGAGAAGG	AAGCCACTGA	CATGGGCAAG	GGTTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTTAAGG	CTGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Strombomonas_napiformis_2	CATCGAGAAG	TTTCGAGAAGG	AGGCCCTCTGA	GATGGGCAAG	GGGTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTTAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Strombomonas_napiformis_gDNA_3	CATCGAGAAG	TTTCGAGAAGG	AGGCCCTCTGA	GATGGGCAAG	GGGTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTTAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[220]
Strombomonas_acuminata_SAG1280_1_2	-----GG	TTTCGAGAAGG	AGGCCGCCGA	GATGGGCAAG	GGGTCCTTCA	AGTATGCCTG	GGTGCTGGAC	AAGCTTAAGG	CCGAGCGTGA	GCAGCGGTATC	ACCATCGATA	[102]

	230	240	250	260	270	280	290	300	310	320	330	
Entosiphon_sulcatum_FJ807254	TTGCCCTGTG	GAAAGTTCGAG	ACCCCAAGT	CCGTGTTTAC	CATCATTGAC	GCCCGGGAC	ACCGTGACTT	CATCAAGAAC	ATGATCACTG	GTACTTCCCA	GGCAGACGCC	[278]
Entosiphon_sp	TTGCCCTGTG	GAAAGTTCGAG	ACCCCAAGT	CCGTGTTTAC	CATCATTGAC	GCCCGGGAC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GTACTTCCCA	GGCAGATGCC	[330]
Ploeotia_oblonga_gDNA_2	TTGCCCTGTG	GAAAGTTCGAG	ACGCCCAAGT	CCGTCTTAC	GATCATTGAT	GCCCCGGAC	ACCGTGACTT	CATCAAGAAC	ATGATCACGG	GAACCTCTCA	GGCCGACGCT	[330]
Peranema_trichophorum_FJ807243_2	TtGcCctTtTg	GAAaTtCgAa	AcTgCcAaGt	CaGtCtTcAc	CaTcAtTgAt	GcTcCaggAc	AtCgTgAcTt	CaTcAagAaC	AtGaTtAcCg	GaAcTtCaCa	AgCtGAtGcC	[324]
Peranema_trichophorum_2	TCGCCCTGTG	GAAAGTTCGAG	ACAGCCAAAT	CAGTCTTAC	CATTATTGAT	GCCCCAGGAC	ACAGAGACTT	CATCAAGAAC	ATGATTACTG	GAACCTTCCA	AGCTGATGCT	[330]
Euglena_stellata_2	TTGCACCTTG	GAAAGTTGAA	ACTGCCAAGT	CCGTCTTAC	TATCATTGAC	GCCCCAGGAC	ATCGTGATTT	TATYAAGAAC	ATGATCACTG	GCACCTCTCA	GGCTGATGCT	[210]
Euglena_hiemalis_1	TTGCCTTTGT	GAAAGTTCGAG	ACTGCAAGT	CTGTCTTAC	CATTATTGAT	GCTCCAGGAC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GAACCTCTCA	GGCTGATGCT	[330]
Euglena_hiemalis_2	TTGCCCTGTG	GAAAGTTCGAA	ACTGCCAGAT	CCGTCTTAC	CATCATTGAT	GCACCTGGCC	ATCGTGATTT	CATCAAGAAC	ATGATTACCG	GCACGTCTCA	GGCCGATGCT	[330]
Euglena_exilis	TCGCCCTGTG	GAAATTCGAA	ACGCCCAAGT	CTGTGTTAC	CATCATTGAT	GCCCCTGCCC	ATCGTGACTT	CATCAAGAAC	ATGATTACTG	GAACCTCCCA	AGCTGATGCA	[210]
Strombomonas_triquetra_1	TTGCCCTGTG	GGAGTTTGAG	ACAGCCAAGT	CTGTGTTAC	CATCATTGAT	GCACCTGGCC	ACCGTGACTT	CATCAAGAAC	ATGATCACAG	GCACCTCTCA	GGCAGATGCA	[330]
Strombomonas_napiformis_1	TTGCCCTGTG	GAAAGTTCGAG	ACYGCCAAGT	CAGTGTTCAC	AATTATTGAT	GCTCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GAACATCCCA	AGCTGATGCA	[330]
Strombomonas_acuminata_ACOI_2467	TTGCCCTGTG	GAAAGTTTGAG	ACTGCCAAGT	CAGTGTTCAC	CATTATTGAT	GCTCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATTACCG	GAACATCCCA	GGCTGATGCA	[330]
Strombomonas_acuminata_SAG1280_1_1	TTGCCCTGTG	GAAAGTTTGAG	ACTGCCAAGT	CAGTGTTCAC	CATTATTGAT	GCTCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACCG	GAACATCCCA	GGCTGACGCA	[212]
Phacus_orbicularis	TTGCCCTGTG	GAAATTCGAG	ACTGCCAAGT	CAGTCTTAC	CGTCATTGAC	GCTCCCGGCC	ATCGTGACTT	CATCAAGAAC	ATGATTACCG	GCACGTCTCA	GGCAGACGCT	[288]
Euglena_stellata_2_2	TTGCTTTGTG	GAAATTTGAA	ACGCCCAAGC	CTGTGTTTAC	TATCATCGAT	GCTCCCGGCC	ATCGTGATTT	CATCAAGAAC	ATGATTACTG	GTACTTCTCA	GGCTGATGCT	[207]
Trachelomonas_similis	TTGCTCTCTG	GAAATTCGAA	ACTGCCAAGT	CAGTCTTAC	CATCATTGAT	GCTCCTGGAC	ACCGAGACTT	CATCAAGAAC	ATGATCACAG	GGACCTCACA	GGCTGACGCT	[210]
Trachelomonas_volvocina	TTGCTCTGTG	GAAAGTTCGAG	ACAGCCAAGT	CAGTCTTAC	GATCATTGAT	GCTCCTGGTC	ACCGAGATTT	CATCAAGAAC	ATGATCACAG	GGACTTCTCA	GGCTGATGCT	[330]
Euglena_clara	TCGCCCTGTG	GAAATTCGAA	ACTGCGAAGT	CCGTGTTAC	CATCATCGAT	GCCCCGGGCC	ACCGCGACTT	CATCAAGAAC	ATGATCACTG	GCACGTCTCA	GGCCGATGCT	[210]
Euglena_deses	TTGCCCTGTG	GAAAGTTCGAA	ACGCCCAAGT	CCGTGTTAC	AATTATTGAT	GCACACGGCC	ATCGTGATTT	CATCAAGAAC	ATGATCACTG	GCACATCCCA	GGCCGATGCT	[330]
Lepocinclis_ovum	TCGCCCTGTG	GAAATTCGAA	ACTGCCAAGT	CCGTCTTAC	CGTTATCGAT	GCCCCTGCCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GCACGTCTCA	GGCAGACGCC	[210]
Euglena_granulata	TCGCCCTCTG	GAAAGTTCGAG	ACAGCTAAGT	CTGTCTTAC	TATCATTGAT	GCTCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATTACTG	GCACATCGCA	GGCTGATGCT	[330]
Trachelomonas_planctonica	TTGCGCTGTG	GAAAGTTCGAG	ACAGCCAAGT	CAGTGTTCAC	AATCATCGAC	GCCCCCGGCC	ATCGTGACTT	TATCAAGAAC	ATGATCACTG	GGACGTCTCA	GGCTGATGCA	[212]
Trachelomonas_lefevrei	TTGCGCTGTG	GAAAGTTCGAG	ACGCCGAGGT	CAGTGTTCAC	AATCATCGAC	GCCCCTGGCC	ATCGTGACTT	TATCAAGAAC	ATGATCACTG	GAACGTCTCA	GGCTGATGCC	[330]
Trachelomonas_echinata	TCGCCCTTTG	GAAATTCGAA	ACTGCAAAAT	CTGTTTTAC	CATCATCGAC	GCCCCCGGAC	ATCGGGATTT	CATCAAGAAC	ATGATTACTG	GCACATCTCA	GGCTGATGCT	[330]
Discoplastis_adunca	TTGCCCTCTG	GAAATTCGAG	ACGCCCAAGT	CTGTCTTAC	CGTGATTGAT	GCCCCCGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GAACCTCTCA	GGCTGACGCT	[330]
Lepocinclis_trieteris_1	TTGCCCTGTG	GAAATTCGAG	ACTGCAAAAT	CAGTGTTCAC	AGTCATTGAC	GCCCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GCACATCTCA	GGCTGATGCT	[330]
Lepocinclis_acus_var_major	TCGCCCTGTG	GAAAGTTCGAG	ACTGCGAAGT	CAGTCTTAC	CATCATTGAT	GCCCCTGGAAC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GCACGTCTCA	GGCAGATGCT	[330]
Lepocinclis_steinii	TTGCCCTGTG	GAAATTCGAA	ACTGCCAAGT	CTGTCTTAC	AATCATCGAT	GACACTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GCACGTCTCA	GGCAGACGCT	[210]
Monomorphina_aenigmatica_2	TTGCGCTGTG	GAAATTCGAG	ACAGCCAAGT	CCGTCTTAC	AATTATCGAT	GCTCCTGGCC	ACCGCGACTT	CATCAAGAAC	ATGATCACAG	GCACGTCTCA	GGCAGATGCT	[330]
Strombomonas_ovalis_2	TTGCGCTGTG	GAAATTCGAG	ACAGCCAAGT	AATTATCGAT	AGTCATTGAC	GCCCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACAG	GCACGTCTCA	GGCAGATGCT	[212]
Strombomonas_borysteniensis_3	TTGCGCTGTG	GAAATTCGAG	ACAGCCAAGT	CCGTCTTAC	AATTATCGAT	GCCCCTGGCC	ACCGCGACTT	CATCAAGAAC	ATGATCACAG	GCACGTCTCA	GGCAGATGCT	[202]
Phacus_pusillus	TTGCCCTGTG	GAAATTCGAG	ACGCCCAAGT	CAGTCTTAC	GGTCATCGAC	GCTCCCGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACGG	GCACGTCTCA	GGCAGACGCC	[330]
Lepocinclis_trieteris_2	TTGCCCTGTG	GAAATTCGAG	ACGCCCAAGT	CAGTCTTAC	GGTCATCGAC	GCTCCCGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACGG	GCACGTCTCA	GGCAGACGCC	[212]
Trachelomonas_conradi	TTGCCCTGTG	GAAAGTTCGAG	ACTGCGAAGT	CAGTTTTAC	AATCATCGAT	GCCCCAGGAC	ATCGTGACTT	CATCAAGAAC	ATGATTACTG	GTACCTCCCA	GGCTGACGCC	[330]
Monomorphina_sp	TCGCTCTGTG	GAAATTCGAG	ACGCCGAAAT	CCGTCTTAC	CATCATTGAT	GCTCCTGGCC	ATCGTGACTT	TATCAAGAAC	ATGATCACTG	GCACCTCCCA	GGCTGACGCT	[330]
Euglena_gracilis_gDNA_3	TCGCCCTGTG	GAAAGTTCGAG	ACAGCCAAGT	CTGTCTTAC	AATTATTGAC	GCTCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATTACTG	GTACCTCTCA	GGCTGACGCT	[330]
Eutreptiella_pomquetensis	TCGCCCTGTG	GAAAGTTCGAG	ACAGCCAAGT	CTGTCTTAC	AATTATTGAC	GCTCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATTACTG	GTACCTCTCA	GGCTGACGCT	[330]
Eutreptia_pertyi	TCGCCCTGTG	GAAAGTTCGAG	ACAGCCAAGT	CTGTCTTAC	AATTATTGAC	GCTCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATTACTG	GTACCTCTCA	GGCTGACGCT	[330]
Monomorphina_pyrum	TCGCCCTGTG	GAAAGTTCGAG	ACAGCCAAGT	CTGTCTTAC	AATTATTGAC	GCTCCTGGCC	ATCGTGACTT	CATCAAGAAC	ATGATTACTG	GTACCTCTCA	GGCTGACGCT	[330]
Cryptoglena_pigra_gDNA_4	TCGCCCTGTG	GAAAGTTCGAA	ACTGCTAAGT	CCGTCTTAC	CATCATTGAT	GCCCCTGGAAC	ACCGTGATTT	CATCAAGAAC	ATGATCATTG	GTACCTCTCA	GGCTGACGCT	[330]
Euglena_gracilis_gDNA_2	TCGCCCTGTG	GAAAGTTCGAA	ACTGCTAAGT	CCGTCTTAC	CATCATTGAT	GCCCCTGGAAC	ACCGTGATTT	CATCAAGAAC	ATGATCATTG	GTACCTCTCA	GGCTGACGCT	[330]
Euglena_gracilis_1	TCGCCCTGTG	GAAAGTTCGAA	ACTGCTAAGT	CCGTCTTAC	CATCATTGAT	GCCCCTGGAAC	ACCGTGATTT	CATCAAGAAC	ATGATCATTG	GTACCTCTCA	GGCTGACGCT	[330]
Ploeotia_oblonga_1	TCGCTCTCTG	GAAAGTTCGAG	ACTGCCAAGT	CTGTTTTAC	CATCATTGAT	GCACCCGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACCG	GTACTTCCCA	GGCCGATGCC	[330]
Euglena_geniculata	TCGCTCTCTG	GAAAGTTCGAG	ACTGCCAAGT	CTGTTTTAC	CATCATTGAT	GCACCCGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACCG	GTACTTCCCA	GGCCGATGCC	[330]
Cryptoglena_pigra_2	YCGCTCTCTG	GAAAGTTCGAG	ACAGCCAAGT	CTGTCTTAC	TATCATCGAT	GCCCCAGGCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GGACATCCCA	GGCTGATGCC	[330]
Cryptoglena_pigra_gDNA_3	TCGCTCTTTG	GAAAGTTCGAG	ACAGCCAAGT	CTGTCTTAC	CATCATCGAT	GCCCCGGGCC	ATCGCGACTT	CATCAAGAAC	ATGATCACTG	GAACGTCTCA	GGCTGATGCA	[330]
Cryptoglena_skujai_2	TCGCTCTTTG	GAAAGTTCGAG	ACAGCCAAGT	CTGTCTTAC	CATCATCGAT	GCCCCGGGCC	ATCGCGACTT	CATCAAGAAC	ATGATCACTG	GAACGTCTCA	GGCTGATGCA	[330]
Cryptoglena_pigra_1	TTGCCCTCTG	GAAATTCGAG	ACTGCAAAAT	CAGTGTTCAC	CATCATTGAT	GCCCCTGCCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GGACATCCCA	GGCTGATGCT	[330]
Cryptoglena_skujai_1	TCGCCCTGTG	GAAATTCGAG	ACTGCAAAAT	CAGTGTTCAC	CATCATTGAT	GCCCCTGCCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GGACATCCCA	GGCTGATGCT	[330]
Colacium_vesiculosum	TCGCCCTCTG	GAAAGTTCGAG	ACGCCCAAGT	CCGTCTTAC	AATCATCGAT	GCCCCTGTTT	ACCGGGACTT	CATCAAGAAC	ATGATCACCG	GGACTTCCCA	GGCTGATGCT	[289]
Colacium_sp	TTGCCCTCTG	GAAAGTTCGAG	ACGCCCAAGT	CTGTGTTAC	CATCATCGAT	GCCCCTGCCC	ATCGTGACTT	CATCAAGAAC	ATGATTACCG	GGACCTCCCA	GGCCGATGCT	[330]
Colacium_elongatum	TTGCCCTCTG	GAAAGTTCGAG	ACTGCAAAAT	CTGTGTTAC	CATCATCGAT	GCCCCTGCCC	ATCGTGACTT	CATCAAGAAC	ATGATTACCG	GGACCTCCCA	GGCCGATGCT	[330]
Strombomonas_triquetra_2	TCGCCCTGTG	GAAAGTTCGAG	ACTGCCAAGT	CTGTGTTAC	CATCATTGAT	GCCCCTGTTT	ACCGGGACTT	CATCAAGAAC	ATGATTACAG	GGACCTCTCA	GGCTGATGCT	[330]
Strombomonas_ovalis_1	TTGCCCTGTG	GAAAGTTCGAG	ACTGCCAAGT	CTGTGTTAC	CATCATTGAT	GCCCCCGGCC	ATCGTGATTT	CATCAAGAAC	ATGATCACTG	GGACTTCTCA	GGCTGATGCT	[330]
Monomorphina_aenigmatica_1	TTGCCCTGTG	GAAAGTTCGAG	ACTGCAAAAT	CTGTGTTAC	CATCATTGAT	GCCCCCGGCC	ATCGTGATTT	CATCAAGAAC	ATGATCACTG	GGACTTCTCA	GGCTGATGCT	[202]
Strombomonas_borysteniensis_1	TTGCCCTGTG	GAAAGTTCGAG	ACTGCCAAGT	CTGTGTTAC	CATCATTGAT	GCCCCTGTTT	ACCGGGACTT	CATCAAGAAC	ATGATCACTG	GGACTTCTCA	GGCTGATGCT	[210]
Strombomonas_borysteniensis_gDNA_2	TTGCCCTGTG	GAAAGTTCGAG	ACTGCCAAGT	CTGTGTTAC	CATCATTGAT	GCCCCTGCCC	ATCGTGACTT	CATCAAGAAC	ATGATCACTG	GGACTTCTCA	GGCTGATGCT	[330]
Strombomonas_napiformis_2	TTGCCCTCTG	GAAAGTTTGAG	ACTGCCAAGT	CTGTCTTAC	CATCATCGAT	GCCCCGGGAC	ATCGTGATTT	CATCAAGAAC	ATGATCACTG	GGACTTCTCA	GGCTGATGCT	[330]
Strombomonas_napiformis_gDNA_3	TTGCCCTCTG	GAAAGTTTGAG	ACTGCCAAGT	CTGTCTTAC	CATCATCGAT	GCCCCGGGAC	ATCGTGATTT	CATCAAGAAC	ATGATCACTG	GGACTTCTCA	GGCTGACGCT	[330]
Strombomonas_acuminata_SAG1280_1_2	TTGCCCTCTG	GAAAGTTCGAG	ACTGCCAAGT	CTGTCTTAC	CATCATCGAT	GCCCCGGTTC	ACCGTGATTT	CATCAAGAAC	ATGATCACGG	GTACCTCTCA	GGCTGACGCT	[212]

	340	350	360	370	380	390	400	410	420	430	440	
Entosiphon_sulcatum_FJ807254	GCCGTCCTCG	TCATCGACTC	CACCACCTGGT	GGTTTCGAGG	CCTGCATCAG	CAAGGATGGC	CAGACCCGCTG	AGCACGCTCT	CCTCGCCTAC	ACCCTGGGTG	TCAAGCAGAT	[388]
Entosiphon_sp	GCTGTCTTGG	TCATTGACTC	AACCACCTGAT	GGTTTCGAGG	CCGGTATCAG	CAAGGATGGT	CAGACGCGCTG	AGCACGCGCT	GCTCGCCTAC	ACCTTGGGTG	TCAAGCAGAT	[440]
Ploeotia_oblonga_gDNA_2	GCCGTGTTGG	TCATCGATTG	CACCACCGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGACGGG	CAGACCCGCTG	AGCACGCTCT	GTTGGCCTTC	ACCTTGGGAG	TGAAGCAAAT	[440]
Peranema_trichophorum_FJ807243_2	GcTgTcTtGg	TcAtTgAcTc	CaCcAcAGgA	GgTtTcGaaG	CtGgTaTtTc	AaAgGaCgGt	CaAaCcCgTg	AaCaCgCcTt	GtTgCcCtAc	AcCtTgGgTg	TcAaGcAaAt	[434]
Peranema_trichophorum_2	GCCGTTTTGG	TCATTGACTC	CACCACAGGA	GGTTTCGAAG	CCGGTATCTC	CAAGGATGGT	CAAACACAGG	AACACGCTCT	CTTGGCCTAC	ACACTTGGTG	TCAAGCAAAT	[440]
Euglena_stellata_2	GCCGTTTTGG	TTATTGATTG	TACCACCTGGT	GGGTTCGAAG	CTGGAATTTC	AAAGGATGGT	CAAAACCCGTG	AACATGCCTT	GCTTGCTTAC	ACCTTGGGAG	TGAAGCAGAT	[320]
Euglena_hiemalis_1	GCCGTGCTGG	TCATCGATTG	CACCACCTGGT	GGGTTCGAAG	CCGGTATCTC	CAAGGATGGG	CAAAACACGG	AACATGCCTC	GCTTGCTATC	ACCCTTGGTG	TCAAGCAAAT	[440]
Euglena_hiemalis_2	GCGGTTTTGG	TCATCGACTC	CACCACCGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGACGGC	CAGACCCGCTG	AGCACGCTCT	GCTTGCTTAC	ACCCTTGGTG	TCAAGCAAAT	[440]
Euglena_exilis	GCTGTGCTGG	TCATTGACTC	CACCACCTGGT	GGGTTCGAAG	CTGGTATCTC	CAAGGATGGG	CAGACCCGCTG	AGCACGCGCT	CCTTGCTATC	ACACTGGGAG	TGAAGCAAAT	[320]
Strombomonas_triquetra_1	GCAGTCTTGG	TCATTGAYTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGG	CAAAACCCGTG	AGCATGCTCT	GCTTGCTATC	ACTCTTGGTG	TGAAGCAAAT	[440]
Strombomonas_napiformis_1	GCYGTCTTGG	TGATTGACTC	CACCACCTGGT	GGCTTTGAAG	CTGGTATTTT	CAAGGATGGG	CAGACCCGCTG	AGCATGCTCT	GTTGGCCTAC	ACGTTGGGTG	TGAAGCAGAT	[440]
Strombomonas_acuminata_ACOI_2467	GCCGTCTTGG	TCATTGACTC	CACCACCTGGT	GGCTTTGAAG	CTGGTATTTT	CAAGGATGGG	CAGACCCGCTG	AGCATGCTCT	GTTGGCCTAC	ACATTGGGTG	TGAAGCAGAT	[440]
Strombomonas_acuminata_SAG1280_1_1	GCCGTCTTGG	TCATTGACTC	CACCACCTGGT	GGCTTTGAAG	CTGGTATTTT	CAAGGATGGG	CAGACCCGCTG	AGCATGCTCT	GTTGGCCTAC	ACATTGGGTG	TGAAGCAGAT	[322]
Phacus_orbicularis	GCTGTGCTGG	TGATTGACTC	GACCACCTGGT	GGTTTTGAGG	CTGGTATCTC	CAAGGATGGT	CAGACGCGCTG	AGCACGCACT	TCTGGCTTAC	ACCCTTGGTG	TCAAGCAGAT	[398]
Euglena_stellata_2_2	GCTGTGCTGG	TTATTGACTC	CACCACCTGGA	GGTTTCGAGG	CTGGTATTTT	TAAAGATGGT	CAGACTCGTG	AGCATGCTTT	GCTTGCTTAC	ACCCCTGGTG	TGAAGCAAAT	[317]
Trachelomonas_similis	GCTGTCTTGG	TCATCGATTG	CACCACCTGGT	GGTTTTGAAG	CTGGTATCTC	CAAGGATGGT	GAGACACGCG	AGCATGCTTT	GCTCGCTTAC	ACACTTGGTG	TGAAGCAAAT	[320]
Trachelomonas_volvocina	GCCGTGTTGG	TCATTGATTG	AACCACCTGGT	GGTTTTGAGG	CTGGTATTTT	CAAGGATGGG	CAGACCCGAG	AACATGCTCT	GCTTGCTTAC	ACTCTTGGTG	TGAAGCAGAT	[440]
Euglena_clara	GCTGTGCTGG	TCATTGACTC	CACCACCGGT	GGCTTCGAGG	CTGGTATTTT	CAAGGATGGG	CAGACTCGTG	AGCATGCCCT	GCTGGCCTAC	ACGCTGGGTG	TAAAGCAAAT	[320]
Euglena_deses	GCAGTCTTGG	TGATTGACTC	CACCACCGGA	GGCTTCGAGG	CTGGCATTTT	CAAGGATGGG	CAGACCCGCTG	AACACGCTCT	GCTTGCTTAC	ACCCTTGGTG	TCAAGCAGAT	[440]
Lepocinclis_ovum	GCCGTTTTGG	TGATCGACTC	CACCACCTGGT	GGTTTTGAAG	CTGGTATCTC	TAAAGATGGA	CAGACGCGTG	AGCACGCTCT	GTTGGCTTAC	ACCTTGGGTG	TTAAGCAGAT	[320]
Euglena_granulata	GCTGTCTTGG	TTATCGACAG	CACCACCTGGT	GGTTTTGAGG	CTGGTATCTC	GAAAGGATGG	CAGACACGCTG	AGCACGCACT	GCTGGCTTAC	ACCCTTGGTG	TGAAGCAGAT	[440]
Trachelomonas_planctonica	GCAGTGTGG	TGATCGACTC	TACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGT	CAGACTCGTG	AACACGCTCT	GCTTGCTTAC	ACACTGGGTG	TGAAGCAGAT	[322]
Trachelomonas_lefevrei	GCAGTGTGG	TGATCGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGG	CAGACTCGTG	AACACGCTCT	GCTTGCTTAC	ACACTGGGTG	TGAAGCAGAT	[440]
Trachelomonas_echinata	GCCGTTTTGG	TCATCGATTG	CACCACCTGGT	GGTTTTGAAG	CAGGTATCTC	CAAGGATGGG	CAAACTCGTG	AACATGCTCT	GCTTGCTTAC	ACATTGGGTG	TGAAGCAGAT	[440]
Discoplastis_adunca	GCTGTCTTGG	TGATTGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGG	CAGACGCGTG	AGCACGCTCT	GCTGGCTTAC	ACACTTGGTG	TGAAGCAGAT	[440]
Lepocinclis_tripteris_1	GCTATCTCTG	TGATCGATTG	AACCACCTGGT	GGTTTCGAGG	CTGGTATTTT	CAAGGATGGG	CAGACGCGTG	AGCACGCTCT	GCTGGCTTAC	ACCTTGGGTG	TGAAGCAGAT	[440]
Lepocinclis_acus_var_major	GCCGTCTCTG	TGATCGACTC	CACCACCTGGT	GGTTTTGAGG	CTGGTATTTT	CAAGGATGGG	CAGACGCGTG	AGCACGCTCT	GCTTGCTTAC	ACTCTGGGTG	TAAAGCAGAT	[440]
Lepocinclis_steinii	GCCGTCTTGG	TGATCGACTC	CACCACCTGGT	GGTTTTGAGG	CTGGTATCTC	CAAGGATGGG	CAGACGCGTG	AGCACGCTCT	GCTGGCTTAC	ACCCTGGGTG	TGAAGCAGAT	[320]
Monomorphina_aenigmatica_2	GCCGTGCTGG	TCATTGACTC	CACCACCTGGT	GGTTTTGAGG	CTGGTATCTC	CAAGGACGGG	CAGACGCGTG	AGCACGCTCT	GCTGGCTTAC	ACCCTTGGTG	TTAAGCAGAT	[440]
Strombomonas_ovalis_2	GCCGTGCTGG	TCATTGACTC	CACCACCTGGT	GGTTTTGAGG	CTGGTATCTC	CAAGGATGGG	CAGACGCGTG	AGCACGCTCT	GCTGGCTTAC	ACCCTTGGTG	TTAAGCAGAT	[322]
Strombomonas_borysteniensis_3	GCCGTGCTGG	TCATTGACTC	CACCACCTGGT	GGTTTTGAGG	CTGGTATCTC	CAAGGATGGG	CAGACGCGTG	AGCACGCTCT	GCTGGCTTAC	ACCCTTGGTG	TTAAGCAGAT	[312]
Phacus_pusillus	GCCGTCTTGG	TCATCGACTC	CACCACCTGGT	GGTTTCGAGG	CAGGTATCTC	CAAGGATGGG	CAGACGCGTG	AGCACGCTCT	GCTGGCTTAC	ACCCTTGGTG	TTAAGCAGAT	[440]
Lepocinclis_tripteris_2	GCCGTCTTGG	TCATCGACTC	GACCACCTGGT	GGTTTCGAGG	CAGGTATCTC	CAAGGATGGG	CAGACGCGTG	AGCACGCTCT	GCTGGCTTAC	ACCCTTGGTG	TTAAGCAGAT	[322]
Trachelomonas_conradi	GCCGTGCTGG	TGATTGATTG	CACCACCTGGT	GGTTTCGAGG	CAGGTATCTC	CAAGGATGGA	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACGCTGGGTG	TGAAGCAGAT	[440]
Monomorphina_sp	GCAGTGTGG	TTATCGACTC	CACCACCTGGT	GGTTTCGAAg	CTGGTATCTC	CAAGGATGGG	CAAAACCCGTG	AGCATGCTCT	TCTGGCCTAC	ACTCTGGGTG	TGAAGCAGAT	[440]
Euglena_gracilis_gDNA_3	GCGGTGCTGG	TTATCGACTC	CACGACTGGT	GGTTTCGAGG	CTGGTATCTC	TAAAGACGGT	CAGACCCGCTG	AGCACGCTCT	TCTGGCCTAC	ACTCTGGGTG	TGAAGCAGAT	[440]
Eutreptiella_pomquetensis	GCGGTGCTGG	TTATCGACTC	CACGACTGGT	GGTTTCGAGG	CTGGTATCTC	TAAGGACGGT	CAGACCCGCTG	AGCACGCTCT	TCTGGCCTAC	ACTCTGGGTG	TGAAGCAGAT	[440]
Eutreptia_pertyi	GCGGTGCTGG	TTATCGACTC	CACGACTGGT	GGTTTCGAGG	CTGGTATCTC	TAAGGACGGT	CAGACCCGCTG	AGCACGCTCT	TCTGGCCTAC	ACTCTGGGTG	TGAAGCAGAT	[440]
Monomorphina_pyrum	GCGGTGCTGG	TTATCGACTC	CACGACTGGT	GGTTTCGAGG	CTGGTATCTC	TAAGGACGGT	CAGACCCGCTG	AGCACGCTCT	TCTGGCCTAC	ACTCTGGGTG	TGAAGCAGAT	[440]
Cryptoglena_pigra_gDNA_4	GCAGTGTGG	TCATTGATTG	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGA	CAGACTCGTG	AGCATGCTCT	GCTGGCTTAC	ACCCTGGGTG	TGAAGCAAAT	[440]
Euglena_gracilis_gDNA_2	GCAGTGTGG	TCATTGATTG	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGA	CAGACTCGTG	AGCATGCTCT	GCTGGCTTAC	ACCCTGGGTG	TGAAGCAAAT	[440]
Euglena_gracilis_1	GCTGTCTTGG	TCATTGACAG	CACAACCTGGT	GGTTTCGAGG	CCGGTATCTC	CAAGGATGGT	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACCCTTGGTG	TAAAGCAGAT	[440]
Ploeotia_oblonga_1	GCTGTCTTGG	TCATTGACAG	CACAACCTGGT	GGTTTCGAAg	CCGGTATCTC	CAAGGATGGT	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACCCTTGGTG	TAAAGCAGAT	[440]
Euglena_geniculata	GCTGTCTTGG	TCATTGACAG	CACAACCTGGT	GGTTTCGAAg	CCGGTATCTC	CAAGGATGGT	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACCCTTGGTG	TAAAGCAGAT	[440]
Cryptoglena_pigra_2	GCAGTGTGG	TGATTGATTG	TACAACCTGGT	GGTTTCGAGG	CTGGTATTTT	CAAGGATGGG	CAGACCCGCTG	AGCACGCTCT	GTTGGCCTAT	ACCCTCGGTG	TGAAGCAGAT	[440]
Cryptoglena_pigra_gDNA_3	GCTGTCTTGG	TGATTGACTC	CACCACCTGGT	GGTTTCGAGG	CCGGTATCCC	CAAGGATGGT	CAGACCCGCTG	AGCACGCGCT	GCTGGCCTAC	ACCCTTGGTG	TGAAGCAGAT	[440]
Cryptoglena_skujai_2	GCTGTCTTGG	TGATTGACTC	CACCACCTGGT	GGTTTCGAGG	CCGGTATCTC	CAAGGATGGT	CAGACCCGCTG	AGCACGCGCT	GCTGGCCTAC	ACCCTTGGTG	TGAAGCAGAT	[440]
Cryptoglena_pigra_1	GCGGTCTTGG	TGATTGACTC	CACCACCTGGT	GGTTTCGAGG	CCGGTATTTT	CAAGGATGGG	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACCCTTGGTG	TGAAGCAGAT	[440]
Cryptoglena_skujai_1	GCCGTGCTGG	TGATTGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATTTT	CAAGGATGGG	CAGACCCGCTG	AGCATGCTCT	GCTGGCCTAC	ACCCTTGGTG	TGAAGCAGAT	[440]
Colacium_vesiculosum	GCCGTGTTGG	TGATTGACTC	CACCACCTGGT	GGCTTTGAGG	CTGGTATCTC	CAAGGACGGT	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACCCTGGGGG	TGAAGCAGAT	[399]
Colacium_sp	GCCGTGCTGG	TGATTGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGG	CAGACCCGCTG	AGCACGCTCT	GTTGGCCTAC	ACCCTGGGTG	TGAAGCAGAT	[440]
Colacium_elongatum	GCCGTGCTGG	TGATTGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGG	CAGACCCGCTG	AGCACGCGCT	GCTGGCCTAC	ACGTTGGGTG	TGAAGCAGAT	[440]
Strombomonas_triquetra_2	GCCGTGCTGG	TCATCGATTG	TACTACTGGT	GGTTTCGAGG	CTGGTATTTT	CAAGGATGGG	CAGACTCGTG	AGCATGCGCT	GCTGGCTTAC	ACGCTGGGTG	TGAAGCAAAT	[440]
Strombomonas_ovalis_1	GCTGTGCTGG	TTATCGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGT	CAGACACGCTG	AGCACGCGCT	GCTGGCCTAC	ACCCTGGGTG	TGAAGCAGAT	[440]
Monomorphina_aenigmatica_1	GCTGTGCTGG	TTATCGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGT	CAGACACGCTG	AGCACGCGCT	GCTGGCCTAC	ACCCTGGGTG	TGAAGCAGAT	[312]
Strombomonas_borysteniensis_1	GCTGTGCTGG	TTATCGACTC	TACCACCTGGT	GGCTTTGAGG	CTGGTATCTC	CAAGGATGGT	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACCCTGGGTG	TGAAGCAAAT	[320]
Strombomonas_borysteniensis_gDNA_2	GCTGTGCTGG	TTATCGACTC	TACCACCTGGT	GGTTTCGAGG	CTGGTATCTC	CAAGGATGGT	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACCCTGGGTG	TGAAGCAAAT	[440]
Strombomonas_napiformis_2	GCTGTGCTGG	TTATCGACTC	CACCACCTGGT	GGTTTCGAGG	CAGGTATTTT	CAAGGATGGT	CAGACCCGCTG	AGCACGCTCT	GCTGGCCTAC	ACTCTGGGTG	TGAAGCAGAT	[440]
Strombomonas_napiformis_gDNA_3	GCTGTGCTGG	TTATCGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATTTT	CAAGGATGGT	CAGACCCGCTG	AGCACGCTCT	GCTTGCTTAC	ACTCTGGGTG	TGAAGCAGAT	[440]
Strombomonas_acuminata_SAG1280_1_2	GCTGTGCTGG	TCATTGACTC	CACCACCTGGT	GGTTTCGAGG	CTGGTATTTT	CAAGGATGGT	CAGACCCGCTG	AGCACGCTCT	GCTTGCTTAC	ACTCTGGGTG	TGAAGCAGAT	[322]

	450	460	470	480	490	500	510	520	530	540	550	
Entosiphon_sulcatum_FJ807254	GATCGTTGCC	TGCAACAAGA	TGGACGACAA	GACCGTCAAG	TACTCCCAGG	CCCGTTATGA	AGAGATCAAG	AAGGAGGTCA	CGGCATACCT	GAAGAAGGTC	GGCTACAACC	[498]
Entosiphon_sp	GATCGTCGCC	TGCAACAAGA	TGGACGACAA	GACCGTCAAG	TACTCCCAGG	CCCGTTATGA	GGAGATCAAG	AAGGAGGTCA	CGCGTACCT	GAAGAAGGTC	GGCTACAACC	[500]
Ploeotia_oblonga_gDNA_2	GATCGTTGCC	TGCAACAAGA	TGGACGACAA	GACCGTCCAG	TACAACAAGG	ATCGTTATGA	GGAGGTCAAG	AAGGAGGTGT	CCGCGTACCT	CAAGAAGGTC	GGCTACAACC	[500]
Peranema_trichophorum_FJ807243_2	GAATGTCGCT	GTCAACAAGA	TGCAACGACA	GAACGTTAAA	TACAACAAG	ACAGATATGA	AGAAATCAAG	AAGGAGGTGT	CTGCTTACCT	CAAGAAAGTT	GGATACAACC	[544]
Peranema_trichophorum_2	GATTGTTGCC	GTTAACAATA	TGGACGACAA	AACAGTTAAA	TACAACAAG	ACAGATATGA	GAAATCAAG	AAGGAGGTCA	GTGCTTACCT	CAAGAAAGTC	GGCTACAACC	[550]
Euglena_stellata_2	GATTGTTGCC	ACCAACAAGT	TTGATGACAA	GACTGTGAAG	TATTCCCAGT	CTCGTTATGA	AGAGATCAAG	AAGGAGGTTT	CTGGGTACCT	CAAGAAGGTT	GGTTACRATC	[430]
Euglena_hiemalis_1	GATCGTTGCA	ACGAACAAGT	TCGATGACAA	AACAGTGAAG	TACTCTCAGT	CACGTTATGA	GAAATCAAG	AAGGAGGTCT	CTGGATATTT	GAAGAAGGTC	GGCTACAATC	[550]
Euglena_hiemalis_2	GATTGTTGCC	ACCAACAAGT	TCGACGACAA	GACCGTAAAG	TACTCCCAGT	CCCGATATGA	GGAGATCAAG	AAGGAGGTCT	CTGGGTACCT	GAAGAAGGTT	GGCTACAACC	[550]
Euglena_exilis	GATCGTTGCC	ACCAACAAGT	TTGATGATAA	AACGGTCAAG	TATTCCACAG	CACGTTATGA	GAAATCAAG	AAGGAGGTGT	CTATGTATTT	GAAGAAGGTC	GGCTACAACC	[430]
Strombomonas_triquetra_1	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACTGTGAAG	TATTCCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTC	GGATACAATC	[550]
Strombomonas_napiformis_1	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	AACGTGTAAA	TATTCCCAGG	CACGATATGA	GGAGATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTT	GGTTACAATC	[550]
Strombomonas_acuminata_ACOI_2467	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	AACGTGTAAA	TATTCCCAGG	CACGATATGA	GGAGATCAAG	AAGGAGGTGT	CTGGTTACTT	GAAGAAGGTT	GGCTACAATC	[550]
Strombomonas_acuminata_SAG1280_1_1	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	AACGTGTAAA	TATTCCCAGG	CACGATATGA	GGAGATCAAG	AAGGAGGTGT	CTGGTTACTT	GAAGAAGGTT	GGCTACAATC	[432]
Phacus_orbicularis	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	AACGTGTAAA	TATTCCCAGG	CACGATATGA	GGAGATCAAG	AAGGAGGTGT	CTGGTTACTT	GAAGAAGGTT	GGCTACAATC	[508]
Euglena_stellata_2_2	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACTGTGAAG	TACTCCCAGT	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTC	GGTTACAACC	[427]
Trachelomonas_similis	GATTGTGGCC	ACCAACAAGT	TCGACGACAA	AACCGTGAAG	TATTCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGGTACCT	GAAGAAGGTT	GGTTACAACC	[430]
Trachelomonas_volvocina	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACAGTCAAG	TACTCCAGG	CCCGTTACGA	AGAAATCAAG	AAGGAGGTGT	CCATGTACTT	GAAGAAGGTT	GGTTACAACC	[550]
Euglena_clara	GATCGTCGCT	ACCAACAAGT	TCGACGACAA	GACAGTGAAG	TACTCTCAGA	CCCGATATGA	GGAGATCAAG	AAGGAGGTGT	CCGGATACCT	CAAGAAGGTT	GGTTACAACC	[430]
Euglena_deses	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACAGTCAAG	TATTCCAGG	CACGATATGA	GAAATCAAG	AAGGAGGTGT	CAGGATATCT	GAAGAAGGTC	GGATACAACC	[550]
Lepocincilis_ovum	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACAGTGAAG	TATTCCAGG	CTCGTTATGA	AGAAATCAAG	AAGGAGGTGT	CCGGCTACTT	GAAGAAGGTT	GGATACAACC	[430]
Euglena_granulata	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACTGTGAAG	TACTCCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTTT	CCACCTACTT	CAAGAAGGTC	GGTTACAACC	[550]
Trachelomonas_planctonica	GATTGTGGCC	ACCAACAAGT	TCGACGACAA	GACCGTGAAG	TACTCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CCACCTACTT	GAAGAAGGTC	GGTTACAACC	[432]
Trachelomonas_ovalis_2	GATTGTGGCC	ACCAACAAGT	TCGACGACAA	GACCGTGAAG	TACTCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CCCTCTACTT	GAAGAAGGTC	GGTTACTACC	[550]
Trachelomonas_echinata	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACCGTCAAG	TACTCCAGG	CTCGTTATGA	AGAAATCAAG	AAGGAGGTTT	CAGGCTACCT	GAAGAAGGTT	GGATACAACC	[550]
Discoplastis_adunca	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GTCTGTGATG	TTCGCTCAGT	CTCGCTATGA	GAAATTAAG	AAGGAGGTGT	CTGGGTACCT	CAAGAAGGTT	GGTTACAACC	[550]
Lepocincilis_trieteris_1	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACCGTGAAG	TATTCCCAGG	CACGTTATGA	GGAGATCAAG	AAGGAGGTGT	CCGGTTACCT	GAAGAAGGTC	GGATACAACC	[550]
Lepocincilis_acus_var_major	GATTGTGGCC	ACCAACAAGT	TTGACGACAA	GACTGTGAAG	TACTCTCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGCCTATCT	GAAGAAGGTC	GGTTACAATC	[550]
Lepocincilis_steinii	GATTGTGGCC	ACCAACAAGT	TTGACGACAA	GACCGTCAAG	TACAATGAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTT	GGATACAACC	[430]
Monomorphina_aenigmatica_2	GATTGTGGCT	ACCAACAAGT	TCGATGACAA	GACCGTGAAG	TACTCCCAGT	CCCGTTATGA	GAAATTAAG	AAGGAGGTGT	CTGGCTATTT	GAAGAAGGTC	GGTTACAACC	[550]
Strombomonas_ovalis_2	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACCGTGAAG	TACTCCCAGT	CCCGTTATGA	GAAATTAAG	AAGGAGGTGT	CAGGCTATTT	GAAGAAGGTC	GGTTACAACC	[432]
Strombomonas_borysteniensis_3	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACCGTGAAG	TACTCCCAGT	CCCGTTATGA	GAAATTAAG	AAGGAGGTGT	CAGGCTATTT	GAAGAAGGTC	GGTTACAACC	[422]
Phacus_puillus	GATTGTGGCC	ACCAACAAGT	TTGACGACAA	GACTGTGAAG	TACTCTCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CCGGATACCT	CAAGAAGGTC	GGCTACAACC	[550]
Lepocincilis_trieteris_2	GATTGTGGCC	ACCAACAAGT	TTGACGACAA	GACTGTGAAG	TACTCTCAGG	CTCGGTATGA	GAAATCAAG	AAGGAGGTGT	CCGGATACCT	CAAGAAGGTC	GGCTACAACC	[432]
Trachelomonas_conradi	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACCGTGAAG	TACTCCCAGG	CCCGCTACGA	AGAAATCAAG	AAGGAGGTGT	CAGGGTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Monomorphina_sp	GATCGTCGCA	ACCAACAAGT	TTGATGACAA	GACTGTGGCC	TACTCTCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACTT	TAAGAAGGTT	GGTTACAATC	[550]
Euglena_gracilis_gDNA_3	GATCGTTGCC	ACCAACAAGT	TTGATGACAA	GACCGTCAAG	TACTCTCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTGT	CCACCTATCT	GAAGAAGGTC	GGCTACAACC	[550]
Eutreptiella_pomquetensis	GATCGTTGCC	ACCAACAAGT	TTGATGACAA	GACCGTCAAG	TACTCTCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTGT	CCACCTATCT	GAAGAAGGTC	GGCTACAACC	[550]
Eutreptia_pertyi	GATCGTTGCC	ACCAACAAGT	TTGATGACAA	GACCGTCAAG	TACTCTCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTGT	CCACCTATCT	GAAGAAGGTC	GGCTACAACC	[550]
Monomorphina_pyrum	GATCGTTGCC	ACCAACAAGT	TTGATGACAA	GACCGTCAAG	TACTCTCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTGT	CCACCTATCT	GAAGAAGGTC	GGCTACAACC	[550]
Cryptoglena_pigra_gDNA_4	GATTGTGGCC	ACTAATAAGT	TCGACGACAA	GACTGTCAAG	TACTCCCAGG	CTCGTATGA	GAAATCAGG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Euglena_gracilis_gDNA_2	GATTGTGGCC	ACTAATAAGT	TCGACGACAA	GACTGTCAAG	TACTCCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Euglena_gracilis_1	GATCGTTGCC	ACTAATAAGT	TCGACGACAA	GACTGTCAAG	TACTCCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGGTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Ploeotia_oblonga_1	GATCGTTGCC	ACCAACAAGT	TCGATGACAA	GACTGTCAAG	TACTCCCAGG	CCCGTTATGA	GGAGATCAAG	AAGGAGGTTT	CGACCTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Euglena_geniculata	GATCGTTGCC	ACCAACAAGT	TCGATGACAA	GACTGTCAAG	TACTCCCAGG	CCCGTTATGA	GGAGATCAAG	AAGGAGGTTT	CGACCTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Cryptoglena_pigra_2	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACTGTCAAG	TACTCCCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Cryptoglena_pigra_gDNA_3	GATAGTGGCC	ACCAACAAGT	TCGATGACAA	GACCGTCAAG	TACTCCCAGG	CACGTTATGA	GGAGATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Cryptoglena_skujai_2	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACCGTCAAG	TACTCCCAGG	CACGTTATGA	GGAGATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Cryptoglena_pigra_1	GATTGTGGCC	ACGAACAAGT	TCGATGACAA	GACTGTGAAG	TACTCTCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACCT	GAAGAAGGTC	GGCTACAACC	[550]
Cryptoglena_skujai_1	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACTGTCAAG	TACTCTCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACTT	GAAGAAGGTC	GGCTACAACC	[550]
Colacium_vesiculosum	GATCGTCGCC	ACCAACAAGT	TCGATGACAA	GACTGTGAAG	TACTCCCAGG	CCCGTTATGA	GGAGATCAAG	AAGGAGGTGT	CTGGATACCT	CAAGAAGGTC	GGCTACAACC	[509]
Colacium_sp	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACTGTGAAG	TACTCCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTTT	CGGATACTT	GAAGAAGGTC	GGCTACAACC	[550]
Colacium_elongatum	GATTGTGGCC	ACCAACAAGT	TTGATGACAA	GACTGTGAAG	TACTCTCAGG	CTCGGTATGA	GAAATTAAG	AAGGAGGTGT	CTGGCTACCT	TAAGAAGGTT	GGTTACAACC	[550]
Strombomonas_triquetra_2	GATCGTTGCC	ACCAACAAGT	TCGATGACAG	GACTGTGAAG	TACTCCCAGG	CACGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGTTACTT	GAAGAAGGTC	GGTTACAACC	[550]
Strombomonas_ovalis_1	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACTGTGAAG	TACTCTCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACCT	CAAGAAGGTC	GGTTACAACC	[550]
Monomorphina_aenigmatica_1	GATTGTGGCC	ACCAACAAGT	TCGATGACAA	GACTGTGAAG	TACTCTCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CTGGCTACCT	CAAGAAGGTC	GGTTACAACC	[422]
Strombomonas_borysteniensis_1	GATTGTGGCT	ACCAACAAGT	TCGATGATAA	GACTGTGAAG	TACTCCCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTTT	CTGGCTACCT	CAAGAAGGTC	GGTTACAACC	[430]
Strombomonas_borysteniensis_gDNA_2	GATTGTGGCT	ACCAACAAGT	TCGATGATAA	GACTGTGAAG	TACTCCCAGG	CCCGTTATGA	GAAATCAAG	AAGGAGGTTT	CTGGCTACCT	CAAGAAGGTC	GGTTACAACC	[550]
Strombomonas_napiformis_2	GATTGTGGCC	ACCAACAAGT	TCGATGATAA	GACAGTGAAG	TACTCCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CGGGCTACTT	GAAGAAGGTC	GGTTACAACC	[550]
Strombomonas_napiformis_gDNA_3	GATTGTGGCC	ACCAACAAGT	TTGATGATAA	GACAGTGAAG	TACTCCCAGG	CTCGTTATGA	GAAATCAAG	AAGGAGGTGT	CGGGCTACTT	GAAGAAGGTC	GGTTACAACC	[550]
Strombomonas_acuminata_SAG1280_1_2	GATTGTGGCA	ACCAACAAGT	TCGATGACAA	GACCGTGAAG	TACTCCCAGG	CTCGATATGA	GAAATCAAG	AAGGAGGTTT	CGGGCTACTT	GAAGAAGGTC	GGTTACAACC	[432]

	560	570	580	590	600	610	620	630	640	650	660	
Entosiphon_sulcatum_FJ807254	CGGACAAGGT	GAACCTTCATC	CCCATCTCCG	GCTGGGTGGG	TGACAACATG	ATTGAGAAGT	CCGACAACAT	G6GCTGGTAC	AAGGGGCTTT	GCCTCATTGA	GGCCCTCGAC	[608]
Entosiphon_sp	CCGACAAGGT	CAACTTCATC	CCCATCTCCG	GCTGGGTGGG	AGACAACATG	ATTGAGAAGT	CCGACAACAT	G6GCTGGTAC	AAGGGGCTTT	GCCTCATTGA	GGCCCTCGAC	[609]
Ploeotia_oblonga_gDNA_2	CCGACAAGGT	CCCCCTCGTG	CCCATCTCCG	GCTGGGTGGG	TGACAACATG	TTGGAGAAGT	CAGACAACAT	G6GATGGTAC	AAGGGGCTTT	CGCTCATCGA	GGCCCTCGAC	[609]
Peranema_trichophorum_FJ807243_2	CcGaRaARGT	McCYTtCaTc	CcAaTcAgTg	gaTgggTtGg	AgAcAaCaTg	AtTgAaGcCa	CcGaRaACAT	GcCaTgGtAc	AaRgGaTcAa	CcCtSaTtGa	YgCcCtCgAc	[654]
Peranema_trichophorum_2	CAGAAAAAGT	ACCCCTTCATC	CCAATCAGTG	GATGGGTGGG	AGACAACATG	ATTGAGAAGT	CCGAAAAACAT	GCCATGGTAC	AAAGGATCAA	CATTGATTGA	TGCCCTCGAC	[660]
Euglena_stellata_2	CTGAGAAGGT	GCCCTTCATC	CCATTTCTG	GCTGGAATGG	GGACAACATG	ATTGAACCTT	CTGAGAACAT	G6GTTGGTAC	AAGGGTTTGA	CACTTATTGG	TGCCCTGGAC	[540]
Euglena_hiemalis_1	CAGAGAAGGT	TCCCTTCATT	CCAATCTCTG	GTTGGAATGG	GGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GATGGTAC	AAGGGGTTGA	CACTCATTGG	AGCTTTGGAC	[660]
Euglena_hiemalis_2	CCGAGAAGGT	CCCCTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGGCTCA	CCCTCATCGG	TGCCCTGGAC	[660]
Euglena_exilis	CTGAAAAAGT	TCCCTTCATC	CCCATCTCTG	GCTGGAACGG	AGACAACATG	ATTGAGGCCCT	CCGAGAATAT	G6GCTGGTAC	AAGGGGCTTA	CTCTCATTGG	TGCACTGGAC	[540]
Strombomonas_triquetra_1	CAGACAAGGT	TCCCTTCATT	CCCATCTCTG	GCTGGAATGG	CGACAATATG	ATCGAGGCCCT	CTGAGAACAT	GCCCTGGTAC	AAGGGGAACA	CTTTGATTGG	TGCCCTGGAT	[660]
Strombomonas_napiformis_1	CTGACAAGGT	GCCCTTCATY	CCCATTTCTG	GCTGGAACGG	TGACAACATG	ATTGAGGCCCT	CTGAGAATAT	GACGTGGTAC	AAGGGGCTGA	CCTTGATTGG	TGCCCTGGAC	[660]
Strombomonas_acuminata_ACOI_2467	CTGACAAGGT	GCCCTTTATC	CCCATTTCTG	GCTGGAACGG	TGACAACATG	ATTGAGGCCCT	CTGAGAATAT	G6GGTGGTAC	AAGGGGCTGA	CCTTGATTGG	TGCCCTGGAT	[660]
Strombomonas_acuminata_SAG1280_1_1	CTGACAAGGT	GCCCTTTATC	CCCATTTCTG	GCTGGAACGG	TGACAACATG	ATTGAGGCCCT	CTGAGAATAT	G6GGTGGTAC	AAGGGGCTGA	CCTTGATTGG	TGCCCTGGAT	[542]
Phacus_orbicularis	CTGACAAGGT	TCCCTTCATC	CCCATCTCTG	GCTGGAACGG	TGACAACATG	ATTGAGGCCCT	TTGACAACAT	GCCCTGGTAC	AAGGGGAACA	CTTTGATTGG	AGCTCTGGAC	[618]
Euglena_stellata_2_2	CCGAAAAAGT	GCCCTTTATT	CCAATTTCTG	GTTGGAATGG	TGACAACATG	ACTGAGCCCT	CTGAGAACAT	G6GTTGGTAC	AAGGGAACCT	CCTTGATTGG	CGCTCTTGAC	[537]
Trachelomonas_similis	CTGAGAAGGT	GCCATTTCATC	CCCATCTCTG	GCTGGAATGG	AGACAACATG	ATTGAGGCCCT	CTGAGAACAT	GAAATGGTAC	AAGGGTTTGA	CCTTGATTGG	TGCACTGGAC	[660]
Trachelomonas_volvocina	CCGACAAGGT	GCCCTTCATC	CCCATTTCTG	GATTGGAATGG	AGACAACATG	ATTGAGGCCCT	CTGAAAACAT	GCCATGGTAC	AAGGGACACA	CCCTCATTGG	TGCTCTGGAC	[660]
Euglena_clara	CCGAGAAGGT	CCCCTTCATC	CCCATCTCCG	GTTGGAATGG	TGATAACATG	ATCGAAGCCTT	CTGAGAACAT	G6GGTGGTAC	AAGGGGCTGA	CCCTCATTGG	TGCCCTCGAC	[540]
Euglena_deses	CTGAGAAAGT	GCCATTTCATC	CCCATCTCTG	GCTGGAACGG	AGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GATGGTAC	AAGGGATTGA	CCTTGATTGG	TGCCCTTGAC	[660]
Lepocinclis_ovum	CAGACAAGGT	TCCATTTCATC	CCCATCTCTG	GATTGGAACGG	TGACAATATG	ATTGATCCCT	CCGACAACAT	GCCCTGGTAC	AAGGGAATGA	CTTTGATCGG	TGCTCTGGAC	[540]
Euglena_granulata	CTGACAAGGT	GCCCTTCATC	CCCATCTCTG	GCTGGAATGG	AGACAACATG	ATTGAGGCCCT	CTGATAATAT	GTCTTGGTAC	AAGGGGAATA	CTCTCATCCA	G6GCTGGAC	[660]
Trachelomonas_planctonica	CTGACAAGGT	GCCATTTCATC	CCCATCTCTG	G6TGGAAATGG	AGACAACATG	ATCGAGGCCCT	CCGAGAACAT	G6GCTGGTAC	AAGGGGCTGA	CCCTGATCGG	TGCTTTGGAC	[542]
Trachelomonas_lefevrei	CTGACAAGGT	GCCATTTCATC	CCCATCTCCG	GATTGGAATGG	TGACAATATG	ATCGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CCCTGATCGG	TGCTTTGGAC	[660]
Trachelomonas_echinata	CTGAGAAGGT	TCCCTTCATC	CCCATCTCTG	GTTGGAATGG	AGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GATGGTAC	AAGGGATTGA	CACTGATTGG	TGCTTTGGAC	[660]
Discoplastis_adunca	CCGACAAGGT	CCCCTTCATC	CCCATCTCTG	G6TGGAAATGG	TGACAACATG	ATCGAGGCCCT	CCGACAATAT	GACTTGGTAC	AAGGGCCCCA	CGCTCATTGG	G6GCTGGAC	[660]
Lepocinclis_trieteris_1	CTGAAAAAGT	GCCCTTCATT	CCCATTTCTG	GTTGGAATGG	GGACAACATG	ATTGAGGCCCT	CTGACAACAT	G6GTTGGTAC	AAGGGACTGA	CCCTGATTGG	TGCTTTGGAC	[660]
Lepocinclis_acus_var_major	CTGACAAGGT	TCCATTTCATC	CCCATCTCTG	GATTGGAATGG	TGACAATATG	ATTGAGGCCCT	CCGAGAACAT	GCCATGGTAT	AAGGGAATGA	CACTGATCGG	TGCCCTGGAT	[660]
Lepocinclis_steinii	CCGAAAAAGT	TCCATTTCATC	CCCATTTCTG	GCTGGAATGG	AGACAACATG	ATTGAGGCCCT	CAGACAACAT	GAAATGGTAC	AAGGGCCCCA	CGCTGATTGG	TGCCCTGGAC	[540]
Monomorphina_aenigmatica_2	CCGAGAAGGT	GCCGTTTCATC	CCCATTTCTG	GCTGGAATGG	AGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATTGG	TGCTCTGGAT	[660]
Strombomonas_ovalis_2	CCGAGAAGGT	GCCGTTTCATC	CCCATTTCTG	GCTGGAACGG	AGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATTGG	TGCTTTGGAT	[542]
Strombomonas_borysteniensis_3	CCGAGAAGGT	GCCGTTTCATC	CCCATTTCTG	GCTGGAACGG	AGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATTGG	TGCTTTGGAT	[532]
Phacus_pusillus	CCGACAAGGT	GCCGTTTCATC	CCCATCTCCG	GCTGGAACGG	TGACAACATG	ATCGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGTCTGA	CGCTCATCGG	TGCTTTGGAT	[660]
Lepocinclis_trieteris_2	CCGACAAGGT	GCCGTTTCATC	CCCATCTCCG	GCTGGAACGG	TGACAACATG	ATTGAGGCCCT	GTGAGAACAT	G6GCTGGTAC	AAGGGTCTGA	CGCTCATCGG	G6GCTTTGGAT	[542]
Trachelomonas_conradi	CCGAGAAGGT	TCCGTTTCATC	CCCATCTCTG	G6TGGAAACGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CCCTCATTGG	G6GCTTTGGAC	[660]
Monomorphina_sp	CTGAGAAGGT	GCCCTTCATT	CCCATCTCTG	G6TGGAAATGG	TGACAACATG	ATCGAGGCCCT	CTGAAAACAT	G6GTTGGTAC	AAGGGGCTGA	CTCTCATTGG	TGCTTTGGAC	[660]
Euglena_gracilis_gDNA_3	CCGACAAGGT	GCCATTTCATC	CCCATCTCTG	GCTGGAACGG	TGATAACATG	ATCGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATCGG	TGCTTTGGAC	[660]
Eutreptiella_pomquetensis	CCGACAAGGT	GCCATTTCATC	CCCATCTCTG	GCTGGAATGG	TGATAACATG	ATCGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATCGG	TGCTTTGGAC	[660]
Eutreptia_pertyi	CCGACAAGGT	GCCATTTCATC	CCCATCTCTG	GCTGGAATGG	TGATAACATG	ATCGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATCGG	TGCTTTGGAC	[660]
Monomorphina_pyrum	CCGACAAGGT	GCCATTTCATC	CCCATCTCTG	GCTGGAATGG	TGATAACATG	ATCGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATCGG	TGCTTTGGAC	[660]
Cryptoglena_pigra_gDNA_4	CCGAGAAGGT	CCCCTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATCGG	G6GCTTTGGAC	[660]
Euglena_gracilis_gDNA_2	CCGAGAAGGT	CCCCTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CTCTCATCGG	G6GCTTTGGAC	[660]
Euglena_gracilis_1	CCGAGAAGGT	CCCCTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CTGATAACAT	G6GCTGGTAC	GAGGGACTGA	CCCTTATTGG	CGCTCTGGAC	[660]
Ploeotia_oblonga_1	CCGACAAGGT	TCCCTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CAGAGAACAT	G6GCTGGTAC	AAGGGCAACA	CCCTGATCCA	G6GCTTTGAT	[660]
Euglena_geniculata	CCGACAAGGT	TCCCTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CAGAGAACAT	G6GCTGGTAC	AAGGGCAACA	CCCTGATCCA	G6GCTTTGAT	[660]
Cryptoglena_pigra_2	CTGACAAGGT	GCCGTTTCATC	CCCATCTCTG	GCTGGAACGG	TGACAACATG	ATTGAGGCCCT	CCGAGAACAT	G6GGTGGTAC	AAGGGAACCT	CGCTCATTGG	TGCCCTGGAC	[660]
Cryptoglena_pigra_gDNA_3	CCGACAAGGT	GCCATTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CCGAGAACAT	GCCCTGGTAC	AAGGGAATGA	CCCTCATTGG	TGCCCTGGAC	[660]
Cryptoglena_skujai_2	CCGACAAGGT	GCCATTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CCGAGAACAT	GCCCTGGTAC	AAGGGAATGA	CCCTCATTGG	TGCCCTGGAC	[660]
Cryptoglena_pigra_1	CTGATAAGGT	GCCGTTTCATC	CCCATTTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CCGAGAACAA	G6GTTGGTAC	AAGGGGCTGA	CTCTGATTGG	TGCTTTGGAC	[660]
Cryptoglena_skujai_1	CTGATAAGGT	GCCATTTCATC	CCCATTTCTG	GCTGGAATGG	TGATAACATG	ATTGAGGCCCT	CTGAGAACAT	G6CTTGGTAC	AAGGGGCTGA	CTCTGATTGG	TGCTTTGGAC	[660]
Colacium_vesiculosum	CCGAGAAGGT	TCCGTTTCATC	CCCATCTCTG	G6TGGAAATGG	TGACAACATG	ATTGATCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGGCTGA	CCCTCATTGG	CGCTTTGGAT	[619]
Colacium_sp	CCGAGAAGGT	GCCCTTCATC	CCCATCTCTG	GATTGGAACGG	TGACAACATG	ATTGAGTCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGACCCA	CCCTCATCGG	TGCCCTGGAC	[660]
Colacium_elongatum	CTGATAAGGT	TCCCTTCATC	CCCATCTCTG	GATTGGAACGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGAACCT	CCCTCATTGG	G6GCTTTGGAC	[660]
Strombomonas_triquetra_2	CTGACAAGGT	GCCATTTCATC	CCCATCTCTG	GCTGGAATGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	GCCCTGGTAC	AAGGGGAACA	CCCTCATTGG	TGCCCTGGAT	[660]
Strombomonas_ovalis_1	CTGACAAGGT	GCCATTTCATC	CCCATCTCTG	GTTGGAATGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGTCTGA	CTCTGATTGG	CGCTCTGGAC	[660]
Monomorphina_aenigmatica_1	CTGACAAGGT	GCCATTTCATC	CCCATCTCTG	GTTGGAATGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGTCTGA	CTCTGATTGG	CGCTCTGGAC	[532]
Strombomonas_borysteniensis_1	CCGACAAGGT	GCCGTTTCATC	CCCATCTCCG	GATTGGAATGG	TGACAACATG	ATTGATCTCT	CTGAGAACAT	G6GCTGGTAC	AAGGGGCTGA	CTTTGATTGG	TGCCCTGGAC	[540]
Strombomonas_borysteniensis_gDNA_2	CCGACAAGGT	GCCGTTTCATC	CCCATCTCCG	GATTGGAATGG	TGACAACATG	ATTGATCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGTCTGA	CTTTGATTGG	TGCCCTGGAC	[660]
Strombomonas_napiformis_2	CCGACAAGGT	CCCATTTCATC	CCCATCTCTG	GATTGGAACGG	GGACAACATG	ATTGAGGCCCT	CTGAGAATAT	GACCTGGTAC	AAGGGAACCT	CCCTGATCGG	TGCCCTGGAC	[660]
Strombomonas_napiformis_gDNA_3	CCGACAAGGT	CCCATTTCATC	CCCATCTCTG	GTTGGAACGG	GGACAACATG	ATTGAGGCCCT	CTGAGAATAT	GACCTGGTAC	AAGGGAACCT	CCCTGATCGG	TGCCCTGGAC	[660]
Strombomonas_acuminata_SAG1280_1_2	CTGACAAGGT	GCCATTTCATC	CCCATCTCTG	GTTGGAATGG	TGACAACATG	ATTGAGGCCCT	CTGAGAACAT	G6GCTGGTAC	AAGGGGCTCA	CCCTGATTGG	TGCCCTGGAT	[542]

	670	680	690	700	710	720	730	740	750	760	770	
Entosiphon_sulcatum_FJ807254	GGCCTGGAGG	CCCCCAAGCG	CCCGAGCGAC	AAGCCGCTTC	GCCTGCCCTC	CCAGGATGTG	TACAAGATCG	CGGGTATCGG	AACCTGTGCC	GTCGGCCGTG	TCGAGACTGG	[718]
Entosiphon_sp	GGTTTGGATG	CCCCGAAGCG	CCCGCTGTGAC	AGCCGCTTCC	GCCTGCCCTC	CCAGGACGTG	TACAAGATTTG	CGGGTATCGG	AACCTGTGCC	GTTGGCCGTG	TCGAGACTGG	[719]
Ploeotia_oblonga_gDNA_2	AACCTCGAGA	CCCCGAAGCG	CCCGCGCGAC	AAGCCCTTCC	GCTTGCCCTT	CGAGGACGTG	TACAAGATCG	CGGGTATCGG	AACGGTCCCC	GTTGGCCGTG	TTGAGACCGG	[720]
Peranema_trichophorum_FJ807243_2	CaATTaGaaC	CaCCTaAaCg	ACCaAaACgAC	AaACCaCTCa	GACTTCaCCT	CCaAGaTgTg	TaCaAaAtTg	GAAGTaTtTg	AaCtGtGcCc	GtCcGgACgTg	TtGaAaACTgg	[764]
Peranema_trichophorum_2	CAATTAGAAC	CACCTAAACG	ACCAAACGAC	AAACCACTCA	GACTTCCACT	CAAGATGTT	TACAAGATTG	GAGGTATTGG	AACCTGTGCC	GTCGGACGTG	TGGAACCTGG	[770]
Euglena_stellata_2	AACCTTGAAG	CACCAAAAGCG	GCCTTCTGAC	AAACCCCTCA	GGCTACCCCT	CCAAGATGTT	TACAAGATTG	GAGGTATTGG	AACCTGTCCA	GTAAGGACGTG	TGGAACCTGG	[650]
Euglena_hiemalis_1	AATTTGGAGC	CCCCAAAGCG	CGCCATCTGAC	AAGCCCTTGC	GTTTGCCCTT	CGAAGATGTT	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGAG	TTGAGACCGG	[770]
Euglena_hiemalis_2	AACCTGGAGC	CCCCAAAGCG	CCCGAGCGAC	AAGCCTCTGC	GCCTGCCCTC	CGAGGATGTC	TACAAGATCG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGTG	TCGAGACCGG	[770]
Euglena_exilis	AACCTTGAAC	CCCCAAAGCG	GCCAAATGAC	AAGCCCTTTC	GACTTCCACT	CCAGGATGTC	TACAAGATTG	GAGGTATCGG	GACTGTGCCCT	GTTGGTCGAG	TCGAGACTGG	[650]
Strombomonas_triquetra_1	GCTTTGGAGC	CACCCAAGCG	TCCTTCAGAC	AAGCCACTGC	GGCTTCCGCT	CGAGGATGTT	TACAAGATTG	GAGGTATCGG	CACCGTGCCA	GTGGGCGGTG	TTGAGACCGG	[770]
Strombomonas_napiformis_1	AACCTTGAAGC	CCCCCAAGCG	GCCCTCCGAC	AAGCCACTGC	GGCTTCCACT	CGAGGATGTT	TACAAGATCG	GAGGTATTGG	CACCTGTGCCA	GTCGGGCGCG	TTGAGACAGG	[770]
Strombomonas_acuminata_ACOI_2467	AACCTTGAAC	CTCCCAAGCG	TCCTTCTGAC	AAGCCACTGC	GGCTTCCGCT	CGAGGACGTG	TACAAGATCG	GAGGTATTGG	CACCTGTGCCA	GTTGGGCGTG	TTGAGACTGG	[770]
Strombomonas_acuminata_SAG1280_1_1	AACCTTGAAC	CCCCCAAGCG	TCCTTCTGAC	AAGCCACTGC	GGCTTCCGCT	CGAGGACGTG	TACAAGATCG	GAGGTATTGG	CACCTGTGCCA	GTTGGGCGTG	TTGAGACTGG	[652]
Phacus_orbicularis	AACCTTGAAGC	CTCCAAAGCG	TCCATCTGAC	AAGCCTCTTC	GCCTGCCCTC	TCAGGATGTC	TACAAGATTG	GAGGTATTGG	AACCTGTGCCA	GTGGGCGGTG	TGGAACCTGG	[728]
Euglena_stellata_2_2	AATCTGGAGC	CACCAAAAGCG	TCCACGCGAC	AAGCCATTGC	GGCTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTTGGTCGTG	TTGAGACAGG	[647]
Trachelomonas_similis	AACCTTGAAGC	CACCAAAAGCG	TCCAAAGTGC	AAGCCCTTGC	GTCTTCCACT	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTTGGGCGTG	TGGAACCTGG	[650]
Trachelomonas_volvocina	AACCTTGAAGC	CTCCCAAGCG	TCCAAGTGC	AAGCCTCTGC	GTCTTCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	AACCTGTGCCCT	GTTGGTCGTG	TGGAACACAGG	[770]
Euglena_clara	AACCTTGAAGC	CCCCCAAGCG	CCCGAAGCGAC	AAGCCACTTC	GGTTGCCCTC	CGAGGATGTT	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTCGGACGTG	TTGAGACCGG	[650]
Euglena_deses	AACCTTGAAGC	CCCCAAAGCG	CCCAAGTGC	AAGCCACTTC	GACTGCCCTC	CGAGGATGTT	TACAAGATCG	GAGGTATTGG	CACGGTGCCT	GTTGGCCGAG	TTGAGACTGG	[770]
Lepocinclis_ovum	GCACCTTGAGC	CTCCGAAGCG	TCCTACGGAC	AAAGCTCTCC	GTCTCCCTCT	TCAGGATGTC	TACAAGATTG	GAGGTATCGG	CACCTGTGCCA	GTCGGCCGTG	TGGAAGACTGG	[650]
Euglena_granulata	TCCTTGGAGC	CTCCAAAGCG	TCCAAATGAC	AAGCCCTTGC	GACTTCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	CACCTGTGCCCT	GTTGGTCGTG	TTGAGACTGG	[770]
Trachelomonas_planctonica	AATTTGGAGC	CCCCGAAGCG	TCCAAGTGC	AAGCCCTTGC	GACTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	AACCTGTGCCCT	GTCGGTCGTG	TGGAAGACTGG	[652]
Trachelomonas_ovalis_2	AATTTGGAGC	CCCCGAAGCG	TCCAAGCGAC	AAGCCCTTGC	GACTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	AACCTGTGCCCT	GTCGGCCGTG	TGGAAGACTGG	[770]
Trachelomonas_echinata	AACCTTGAAGC	CACCAAGCG	GCCAAATGAC	AAACCTCTGC	GTCTTCCCTC	CGAGGATGTT	TACAAGATTG	GAGGTATTGG	AACCTGTGCCA	GTCGGCCGTG	TGGAACCTGG	[770]
Discoplastis_adunca	ATGCTGAGCC	CCCCAAAGCG	TCCCTCTGAC	AAGCCTCTGC	GGCTGCCCTC	CGAGGATGTC	TACAAGATTG	CGGGTATTGG	AACCTGTGCCCT	GTCGGGCGTG	TCGAGACTGG	[770]
Lepocinclis_triapteris_1	AATCTGGAGC	CACCCAAGCG	CCCGCTCTGAC	AAGCCTCTGC	GATTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	GACAGTGCCCT	GTTGGCCGTG	TCGAGACCGG	[770]
Lepocinclis_acus_var_major	GCCTTGGAGC	CCCTTAAGCG	ACCTACCGAC	AAGCCTCTGC	GTCTCCCTCT	TCAGGACGTG	TACAAGATTG	GAGGTATCGG	CACAGTGCCCT	GTTGGTCGTG	TTGAGACTGG	[770]
Lepocinclis_steinii	AATTTGGAGC	CCCCAAAGCG	ACCTACCGAC	AAGCCACTTC	GTCTCCCTCT	TCAGGACGTG	TACAAGATCG	GAGGTATCGG	CACAGTGCCCT	GTTGGCCGTG	TGGAAGACTGG	[650]
Monomorpha_aenigmatica_2	GCTCTGGAGC	CACCAAAAGCG	TCCCAATGAC	AAACCTCTTC	GTCTGCCCTC	GCAAGATGTC	TACAAGATCG	GAGGTATTGG	GACAGTGCCCT	GTCGGCCGTG	TTGAGACAGG	[770]
Strombomonas_ovalis_2	GCCCTGGAAC	CACCAAAAGCG	TCCCAATGAC	AAACCTCTTC	GTCTGCCCTC	GCAAGACGTG	TACAAGATCG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGTG	TGGAAGACAGG	[652]
Strombomonas_borysteniensis_3	GCCCTGGAAC	CACCAAAAGCG	TCCCAATGAC	AAACCTCTTC	GTCTGCCCTC	GCAAGACGTG	TACAAGATCG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGTG	TTGAGACAGG	[642]
Phacus_pusillus	AACCTTGAAGC	CCCCGAAGCG	TCCCAATGAC	AAGCCCTTGC	GGCTGCCCTC	CGAGGACGTG	TACAAGATCG	CGGGTATCGG	GACAGTGCCCT	GTCGGCCGTG	TCGAGACCGG	[770]
Lepocinclis_triapteris_2	AACCTTGAAGC	CCCCGAAGCG	TCCCAATGAC	AAGCCCTTGC	GGCTGCCCTC	CGAGGACGTG	TACAAGATCG	CGGGTATCGG	GACAGTGCCCT	GTCGGCCGTG	TCGAGACCGG	[652]
Trachelomonas_conradi	AACCTTGAAGC	CCCCGAAGCG	CCCGAGCGAC	AAGCCACTGC	GTCTTCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	AACAGTGCCCT	GTTGGTCGTG	TGGAAGACCGG	[770]
Monomorpha_sp	AACCTTGAAC	CGCCAARGCG	CCCGACTGAC	AAGCCTCTGC	GTCTGCCCTC	CGAGGACGTG	TACAAGATTG	GAGGTATTGG	CACCTGTGCCCT	GTTGGTCGTG	TGGAAGACTGG	[770]
Euglena_gracilis_gDNA_3	AACCTTGAAGC	CGCCTAAGCG	CCCGACTGAC	AAGCCACTTC	GGCTGCCCTT	CGAGGATGTC	TACAAGATTG	GAGGTATCGG	CACCTGTGCCCT	GTTGGCCGTG	TCGAGACAGG	[770]
Eutreptiella_pomquetensis	AACCTTGAAGC	CGCCTAAGCG	CCCGACTGAC	AAGCCACTTC	GGCTGCCCTT	CGAGGATGTC	TACAAGATTG	GAGGTATCGG	CACCTGTGCCCT	GTTGGCCGTG	TCGAGACAGG	[770]
Eutreptia_pertyi	AACCTTGAAGC	CGCCTAAGCG	CCCGACTGAC	AAGCCACTTC	GGCTGCCCTT	CGAGGATGTC	TACAAGATTG	GAGGTATCGG	CACCTGTGCCCT	GTTGGCCGTG	TCGAGACAGG	[770]
Monomorpha_pyrum	AACCTTGAAGC	CGCCTAAGCG	CCCGACTGAC	AAGCCACTTC	GGCTGCCCTT	CGAGGATGTC	TACAAGATTG	GAGGTATCGG	CACCTGTGCCCT	GTTGGCCGTG	TCGAGACAGG	[770]
Cryptoglena_pigra_gDNA_4	AACCTTGAAGC	CGCCTAAGCG	CCCGACTGAC	AAGCCACTTC	GGCTGCCCTT	CGAGGATGTC	TACAAGATTG	GAGGTATCGG	CACCTGTGCCCT	GTTGGCCGTG	TCGAGACAGG	[770]
Euglena_gracilis_gDNA_2	AACCTTGAAGC	CACCAAAAGCG	ACCAAGCAAC	AAGCCCTTCC	GCCTGCCCTC	ACAGGATGTC	TACAAGATTG	GAGGTATTGG	AACCTGTGCCCT	GTTGGTCATG	TTGAGACCGG	[770]
Euglena_gracilis_1	AACCTTGAAGC	CACCAAAAGCG	ACCAAGCGAC	AAGCCCTTCC	GCCTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	AACCTGTGCCCT	GTTGGTCGTG	TTGAGACCGG	[770]
Ploeotia_oblonga_1	GCCTTGGAGC	CCCCCAAGCG	ACCCAATGAC	AAGCCCTTCC	GACTGCCCTC	CCAGGACGTG	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGTG	TGGAAGACTGG	[770]
Euglena_geniculata	GCCTTGGAGC	CCCCCAAGCG	ACCCAATGAC	AAGCCCTTCC	GACTGCCCTC	CCAGGACGTG	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGTG	TGGAAGACTGG	[770]
Cryptoglena_pigra_2	GCTTTGGAGG	CACCGAAGCG	CCCGAGTGC	AAGCCTCTGC	GCCTGCCCTC	CGAGGACGTG	TACAAGATCG	GAGGTATTGG	CACCGTGCCCT	GTTGGTCGTG	TGGAAGACCGG	[770]
Cryptoglena_pigra_gDNA_3	GCTTTGGAGC	CTCCCAAGCG	TCCGAGTGC	AAGCCTCTGC	GCCTGCCCTC	CGAGGATGTT	TACAAGATTG	GAGGTATTGG	CACCGTGCCCT	GTTGGTCGTG	TGGAAGACTGG	[770]
Cryptoglena_skujai_2	GCTTTGGAGC	CTCCCAAGCG	TCCGAGTGC	AAGCCTCTGC	GCCTGCCCTC	CGAGGACGTG	TACAAGATTG	GAGGTATTGG	CACCGTGCCCT	GTTGGTCGTG	TGGAAGACTGG	[770]
Cryptoglena_pigra_1	GCCTTGGAGG	CCCCCAAGCG	TCCCAAGCGAC	AAGCCCTTCC	GGCTGCCCTC	CGAGGATGTT	TACAAGATTG	GAGGTATCGG	GACCGTGCCCT	GTGGGCGGAG	TGGAAGACTGG	[770]
Cryptoglena_skujai_1	GCCTTGGAGC	CTCCCAAGCG	CCCGAGCGAC	AAGCCACTGC	GCCTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	CACCTGTGCCCT	GTGGGCGGTG	TGGAAGACTGG	[770]
Colacium_vesiculosum	AACCTTGAAGC	CCCCCAAGCG	CCCAAGCGAC	AAGCCCTTCC	GCCTGCCCTC	CCAGGATGTT	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGAG	TTGAGACCGG	[729]
Colacium_sp	AACCTTGAAGC	CCCCCAAGCG	CCCAAGCGAC	AAGCCCTTGC	GTCTGCCCTC	CCAGGATGTC	TACAAGATCG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGTG	TTGAGACTGG	[770]
Colacium_elongatum	AACCTTGAAGC	CTCCCAAGCG	TCCCAAGCGAC	AAGCCTCTGC	GTCTGCCCTC	CCAGGATGTC	TACAAGATTG	GAGGTATTGG	AACCTGTGCCCT	GTCGGCCGTG	TGGAAGACTGG	[770]
Strombomonas_triquetra_2	GCCCTGGAGC	CACCAAGCG	CCCGAGCGAC	AAGCCACTGC	GCTTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATCGG	AACCTGTGCCCT	GTGGGCGGAG	TGGAAGACTGG	[770]
Strombomonas_ovalis_1	AACCTTGAAGC	CACCAAGCG	CCCAAGCGAC	AAGCCTCTGC	GTCTGCCCTC	CGAGGACGTG	TACAAGATTG	GAGGTATCGG	GACTGTGCCCT	GTCGGCCGAG	TGGAAGACAGG	[770]
Monomorpha_aenigmatica_1	AACCTTGAAGC	CACCAAGCG	CCCAAGCGAC	AAGCCTCTGC	GTCTGCCCTC	CGAGGACGTG	TACAAGATTG	GAGGTATCGG	GACTGTGCCCT	GTCGGCCGAG	TGGAAGACAGG	[642]
Strombomonas_borysteniensis_1	AACCTTGAAGC	CACCAAGCG	CCCGAGCGAC	AAGCCTCTGC	GCCTGCCCTC	CGAGGACGTG	TACAAGATTG	GAGGTATCGG	GACTGTGCCCT	GTCGGCCGAG	TGGAAGACAGG	[650]
Strombomonas_borysteniensis_gDNA_2	AACCTTGAAGC	CACCAAGCG	CCCGAGCGAC	AAGCCTCTGC	GCCTGCCCTC	CGAGGACGTG	TACAAGATTG	GAGGTATCGG	GACTGTGCCCT	GTCGGCCGAG	TGGAAGACAGG	[770]
Strombomonas_napiformis_2	AACCTTGAAGC	CACCAAGCG	CCCAAGCGAC	AAGCCCTTCC	GTCTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGAG	TGGAAGACTGG	[770]
Strombomonas_napiformis_gDNA_3	AACCTTGAAGC	CACCAAGCG	GCCAAATGAC	AAGCCCTTCC	GTCTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATTGG	GACTGTGCCCT	GTCGGCCGAG	TGGAAGACTGG	[770]
Strombomonas_acuminata_SAG1280_1_2	AACCTTGAAGC	CACCTAAGCG	GCCAAAGTGC	AAGCCCTTCC	GTCTGCCCTC	CGAGGATGTC	TACAAGATTG	GAGGTATCGG	GACTGTGCCCT	GTCGGCCGAG	TGGAAGACTGG	[652]

	780	790	800	810	820	830	840	850	860	870	880	
Entosiphon_sulcatum_FJ807254	TATCCTTAAG	CCCGGCGATG	TTGTGACCTT	CGCCCCGGTC	AACCTTGACCA	CTGAGGTCAA	GTCCATCGAG	ATGCACCACG	AGGCCCTGCC	AGAGGCCGTG	CCG6GTGACA	[828]
Entosiphon_sp	TATCCTGAAG	CCCGGTGATG	TCGTCACCTT	CGCCCTCACT	AACCTTGACCA	CTGAGGTCAA	GTCCATTGAG	ATGCACCACG	AGGCCCTGCC	TGAGGCCGTG	CCG6GTGACA	[880]
Ploeotia_oblonga_gDNA_2	AGTGATGAAG	CCTGGTGACG	TCGTCACATT	CGCCCCCAAC	GCTTTGACCA	CTGAGGTCAA	GTCCATTGAG	ATGCATCACG	AGTCCCTCCC	C6AGGCCGTG	CCC6GCGACA	[880]
Peranema_trichophorum_FJ807243_2	AaTcCtCaAG	CCAGGaGAcGy	TtGcTcAcCtT	TgCaCcAaAc	AaCcTaAcCa	CtGaAgTtTaA	GtCcAtCgAa	AtGcAcCaCg	AaGcTcTtGc	CgAaGcCaCt	CcTgGtGAcA	[874]
Peranema_trichophorum_2	AATCCTCAAG	CCAGGAGACG	TTGTCACTTT	TGCACCAAC	AACCTTGACCA	CTGAAGTCAA	GTCCATTGAA	ATGCACCACG	AGGCTCTTGC	CGAAGCCATT	CCTGGTGACA	[880]
Euglena_stellata_2	TGTGCTCAAG	CCAGGTGATG	TTGTCACTTT	TGCTCCAAAT	AACCTTGACCA	CTGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGTCTCTTCT	GGAAGCCATT	CCAGGAGACA	[760]
Euglena_hiemalis_1	GGTTCTGAAG	CCTGGTGACG	TGGTGACCTT	TGCCCCAAC	AACCTTGACCA	CCGAAGTCAA	GTCAAGTCAA	ATGCATCACG	AGGCCGTCAC	YGAGGCCAYC	CCTGGTGACA	[880]
Euglena_hiemalis_2	TGTCTTGAAG	CCTGGTGACG	YGGTGACATT	TGCCCCAAC	AACCTTGACCA	CCGAAGTCAA	GTCCGTGGAG	ATGCACCACG	AGGCTCTCAC	TGAGGCTATT	CCTGGTGACA	[880]
Euglena_exilis	AAATCTCAAG	CCAGGTGATG	TGGTGACCTT	CGCTCCAAT	AACCTTGACCA	CCGAAGTCAA	GTCAAGTCAA	ATGCATCACG	AAGCCTTGGC	TGAAGCTATT	CCAGGAGACA	[760]
Strombomonas_triquetra_1	GGTTTTGAAG	CCTGGTGATG	TGGTGACCTT	TGCTCCCAAC	AATTTTRACCA	CTGAAGTCAA	GTCTGTTGAA	ATGCACCACG	AGGCTTTGCC	GGAGGCTGTG	CCTGGTGACA	[880]
Strombomonas_napiformis_1	AGTGTTGAAG	CCCGGTGACG	TTGTGACTTT	CGCTCTCAAC	AATTTTGACCA	CTGAAGTCAA	GTCTGTTGAA	ATGCATCACG	AGTCGCTGGC	AGAGGCTGTG	CCTGGTGACA	[880]
Strombomonas_acuminata_ACOI_2467	AGTGTTGAAA	CCCGGTGATG	TTGTGACTTT	CGCTCCCAAC	AATTTTGACCA	CTGAAGTCAA	GTCTGTTGAA	ATGCATCACG	AGTCGCTGGC	AGAGGCTGTG	CCTGGTGATA	[880]
Strombomonas_acuminata_SAG1280_1_1	AGTGTTGAAA	CCCGGTGATG	TTGTGACTTT	CGCTCCCAAC	AATTTTGACCA	CTGAAGTCAA	GTCTGTTGAA	ATGCATCACG	AGTCGCTGGC	AGAGGCTGTG	CCTGGTGATA	[762]
Phacus_orbicularis	TGTACTCAAG	CCTGGTGACG	TCGTGACGTT	TGCTCCCAAC	AATCTGACAA	CCGAGGTGAA	GTCTGTTGAG	ATGCACCACG	AAGCTCTGAC	TGAGGCCATT	CCTGGGGACA	[838]
Euglena_stellata_2_2	TGTCCTGAAG	CCTGGTGACG	TGGTTACATT	TGCACCAAC	GATTTGACTA	CTGAGGTCAA	GTCTGTTGAA	ATGCACCACG	AGTCATTGCT	TGAGGCCATT	CCAGGGGACA	[757]
Trachelomonas_similis	TGTCCTGAAA	CCTGGTGATG	TGGTGACCTT	TGCACCAAC	AACCTTGACAA	CCGAGGTGAA	GTCTGTTGAA	ATGCACCACG	AATCTTTGCT	TGAGGCTGTC	CCTGGAGACA	[760]
Trachelomonas_volvocina	TGTTCTAAAG	CCTGGTGACG	TGGTGACCTT	CGCCCCAAT	GCCTTGACAA	CTGAAGTCAA	GTCCGTGAG	ATGCACCACG	AGGCATTGTC	AGAGGCTGTC	CCTGGTGACA	[880]
Euglena_clara	TATTTTGAAG	CCCGGTGATG	TCGTGACATT	TGCGCCGAAC	AACCTTGACAA	CCGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGTCCTTGCT	GGAGGCCATC	CCTGGTGACA	[760]
Euglena_deses	TGTGCTGAAG	CCTGGTGACG	TTGTGACCTT	TGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	ATCAGTCTGAG	ATGCACCACG	AGTCTCTTCT	TGAGGCCATC	CCTGGTGACA	[880]
Lepocinclis_ovum	TGTCCTGAAG	CCTGGTGACG	TTGTGACTTT	TGCCCCGAAC	AACCTTGACCA	CTGAGGTGAA	ATCAGTCTGAG	ATGCACCACG	AGTCACTGCC	AGAAAGCTGT	CCTGGTGACA	[760]
Euglena_granulata	TGTTCTCAAG	CCTGGTGATG	TTGTACCTTT	CGCCCCAAC	AACCTTGACCA	CCGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCCCTGCT	TGAGGCCATC	CCTGGTGACA	[880]
Trachelomonas_planctonica	TGTTTTGAAG	CCTGGGTGATG	TGGTGACCTT	CGCACCACAC	AATCTGACAA	CCGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AAGCCTTGAC	CGAGGCTATT	CCTGGAGACA	[762]
Trachelomonas_lefevrei	TGTCCTGAAG	CCTGGTGACG	TGGTGACCTT	TGCACCAAC	AATCTGACAA	CCGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AAGCCTTGAC	TGAGGCTATT	CCTGGTGACA	[880]
Trachelomonas_echinata	TGTTTTGAAG	CCCGGTGATG	TCGTGACCTT	TGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCTGTTGAG	ATGCACCATG	AGTCTTTGGC	CGAGGCCATC	CCAGGTGACA	[880]
Discoplastis_adunca	GGTGCTGAAG	CCTGGTGATG	TTGTGACCTT	CGCCCCAAC	AACCTTGACCA	CCGAGGTGAA	GTCTGTCGAA	ATGCATCACG	AGTCTCTCAC	TGAGGCCATC	CCTGGTGACA	[880]
Lepocinclis_trieteris_1	TGTCCTGAAG	CCCGGTGACG	TTGTACCTTT	TGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCTGTTGAG	ATGCACCACG	AGGCCCTGAC	CGAGGCTATC	CCAGGTGACA	[880]
Lepocinclis_acus_var_major	TGTTCTGAAG	CCTGGTGATG	TTGTGACCTT	TGCTCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCTGTTGAG	ATGCACCACG	AGTCTCTTGC	AGAGGCTGTG	CCTGGCGACA	[880]
Lepocinclis_steinii	TGTGCTGAAG	CCCGGTGACG	TTGTCACTTT	CGCGCCAAAC	AATTTTGACCA	CCGAGGTGAA	TGCTGTTGAG	ATGCACCACG	AGTCTGTTGC	TGAGGCCATC	CCG6GCGACA	[760]
Monomorphina_aenigmatica_2	GGTGTTGAAG	CCAGGTGACG	TTGTGACGTT	CGCACCACAC	AACCTTGACAA	CTGAGGTGAA	GTCAAGTCAA	ATGCACCACG	AGTCTTTGGC	TGAGGCCATC	CCTGGTGACA	[880]
Strombomonas_borysteniensis_3	GGTGTTGAAG	CCAGGTGATG	TTGTGACCTT	CGCACCACAC	AACCTTGACCA	CTGAGGTGAA	GTCAAGTCAA	ATGCACCACG	AGTCTTTGGC	TGAGGCCATC	CCTGGTGACA	[762]
Phacus_puillus	TGTCCTGAAG	CCTGGTGATG	TTGTGACCTT	CGCCCCAAC	AACCTTGACCA	CTGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGTCTCTGCC	TGAGGCCGTG	CCG6GCGACA	[880]
Lepocinclis_trieteris_2	TGTCCTGAAG	CCTGGTGATG	TTGTGACCTT	CGCCCCAAC	AACCTTGACCA	CTGAAGTCAA	GTCCGTGCAA	ATGCACCACG	AGTCTCTGCC	TGAGGCCGTG	CCC6GCGACA	[762]
Trachelomonas_conradi	TGTCCTGAAG	CCTGGTGATG	TGGTGACCTT	CGCGCCCAAC	AACCTTGACCA	CTGAAGTCAA	GTCCGTGCAA	ATGCACCACG	AGTCTCTGCC	TGAGGCCATC	CCG6GTGACA	[880]
Monomorphina_sp	TGTCCTGAAG	CCTGGTGACG	TTGTGACCTT	TGCCCCAAC	AATCTGACCA	CTGAAGTCAA	GTCCGTGCAA	ATGCACCACG	AGGCTCTGAC	TGAGGCTGTG	CCTGGTGATA	[880]
Euglena_gracilis_gDNA_3	TGTGCTGAAG	CCTGGTGACG	TTGTGACCTT	CGCCCCAAC	AATCTGACCA	CTGAGGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTCAC	TGAGGCCGTG	CCTGGGACA	[880]
Eutreptiella_pomquetensis	TGTGCTGAAG	CCTGGTGACG	TTGTGACCTT	CGCCCCAAC	AATCTGACCA	CTGAGGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTCAC	YGAGGCCGTG	CCTGGGACA	[880]
Eutreptia_pertyi	TGTGCTGAAG	CCTGGTGACG	TTGTGACCTT	CGCCCCAAC	AATCTGACCA	CTGAGGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTCAC	TGAGGCCGTG	CCTGGGACA	[880]
Monomorphina_pyrum	TGTGCTGAAG	CCTGGTGATG	TTGTGACCTT	CGCCCCAAC	AATCTGACCA	CTGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTCAC	TGAGGCCGTG	CCTGGGACA	[880]
Cryptoglena_pigra_gDNA_4	TGTGCTGAAG	CCTGGTGACG	TTGTGACCTT	CGCCCCAAC	AATCTGACCA	CTGAGGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTCAC	TGAGGCCGTG	CCC6GCGACA	[880]
Euglena_gracilis_gDNA_2	TGTCCTGAAG	CCCGGTGATG	TGGTGACCTT	CGCGCCCAAC	AACCTTGACCA	CTGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTGAC	CGAGGCTGTC	CCTGGCGATA	[880]
Euglena_gracilis_1	TGTCCTGAAG	CCCGGTGATG	TGGTGACCTT	CGCGCCCAAC	AACCTTGACCA	CTGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTGAC	CGAGGCAATC	CCTGGGACA	[880]
Ploeotia_oblonga_1	TGTCCTGAAG	CCCGGTGATG	TTGTGACCTT	CGCCCCAAC	AACCTTGACCA	CTGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTGAC	TGAGGCCATC	CCC6GTGACA	[880]
Euglena_geniculata	TGTCCTGAAG	CCCGGTGATG	TTGTGACCTT	CGCCCCAAC	AACCTTGACCA	CTGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTGAC	TGAGGCCATC	CCC6GTGACA	[880]
Cryptoglena_pigra_2	TGTTTTGAAG	CCGGGTGACG	TTGTGACGTT	TGCGCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGGCCCTGGT	GGAGGCCGTG	CCTGGTGACA	[880]
Cryptoglena_pigra_gDNA_3	TGTTCTGAAG	CCGGGTGACG	TTGTGACGTT	TGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGGCCCTGGC	TGAGGCCGTG	CCTGGTGACA	[880]
Cryptoglena_skujai_2	TGTTCTGAAG	CCGGGTGACG	TTGTGACGTT	TGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGGCCCTGGC	TGAGGCCGTG	CCTGGTGACA	[880]
Cryptoglena_pigra_1	TGTCCTGAAG	CCTGGTGATG	TTGGGACCTT	TGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCAAGTCAA	ATGCACCACG	AGGCTCAGGT	TGAGGCTGTC	CCTGGCGACA	[880]
Cryptoglena_skujai_1	TGTTCTGAAG	CCTGGTGATG	TCGTGACCTT	TGCGCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGGCTTTGGC	AGAGGCCGTG	CCTGGTGACA	[880]
Colacium_vesiculosum	TGTGTTGAAG	CCCGGTGACG	TCGTGACCTT	TGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCTGTTGAG	ATGCACCACG	AGTCCCTGCT	GGAGGCCATC	CCC6GTGACA	[839]
Colacium_sp	TATCTTGAAG	CCTGGTGATG	TCGTGACCTT	CGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGTCTCTTGC	CGAGGCTATC	CCTGGTGACA	[880]
Colacium_elongatum	GGTGTTGAAG	CCTGGTGATG	TTGTGACCTT	CGCCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCTGTTGAG	ATGCACCACG	AGTCTCTTGC	CGAGGCTGTC	CCTGGTGATA	[880]
Strombomonas_triquetra_2	TGTTCTGAAG	CCTGGTGATG	TTGTACCTTT	CGCACCACAC	AACCTTGACCA	CCGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGGCTCTG--	-----	-----	[858]
Strombomonas_ovalis_1	TGTCCTGAAG	CCTGGTGATG	TTGTTACCTT	TGCGCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGTCCCTGCC	TGAGGCTGTC	CCTGGTGACA	[880]
Monomorphina_aenigmatica_1	TGTCCTGAAG	CCTGGTGATG	TTGTTACCTT	TGCGCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGTCCCTGCC	TGAGGCTGTC	CCTGGTGACA	[752]
Strombomonas_borysteniensis_1	TGTCCTGAAG	CCCGGTGATG	TTGTTACCTT	CGCGCCCAAC	AACCTTGACCA	CTGAGGTGAA	GTCCGTGAG	ATGCACCACG	AGTCCCTGCC	TGAGGCTGTC	CCTGGTGACA	[760]
Strombomonas_borysteniensis_gDNA_2	TGTCCTGAAG	CCCGGTGATG	TTGTGACCTT	CGCTCTCAAC	AACCTTGACCA	CTGAGGTGAA	GTCTGTTGAG	ATGCACCACG	AGTCCCTGCC	TGAGGCTGTC	CCC6GCGACA	[880]
Strombomonas_napiformis_2	TGTAATTGAAG	CCTGGTGATG	TCGTGACCTT	TGCACCAAC	AACCTTGACCA	CTGAGGTCAA	GTCTGTCGAA	ATGCACCACG	AGTCATTGGC	TGAGGCTGTC	CCTGGTGACA	[880]
Strombomonas_napiformis_gDNA_3	TGTAATTGAAG	CCTGGTGATG	TCGTGACCTT	TGCACCAAC	AACCTTGACCA	CTGAGGTCAA	GTCTGTCGAA	ATGCACCACG	AGTCATTGGC	TGAGGCTGTC	CCTGGTGACA	[880]
Strombomonas_acuminata_SAG1280_1_2	TGTTCTGAAG	CCCGGTGACG	TCGTGACCTT	TGCACCAAC	AACCTTGACCA	CTGAAGTCAA	GTCTGTTGAG	ATGCACCACG	AGTCTTGGC	TGAGGCCATC	CCTGGGACA	[762]

	890	900	910	920	930	940	950	960	970	980	990	
Entosiphon_sulcatum_FJ807254	ACGTTGGCTT	CAACGTCGAA	AACGTGTCCG	TGAAGGAAAT	TCGCCGTGGC	TTTCGTGCCT	CGAACGCAAA	GAACGACCCC	GCCAAGGAGA	CTGAGACCTT	CACGTGCTCAA	[938]
Entosiphon_sp	ACGTTGGCTT	CAACGTGAA	AACGTGTCCG	TGAAGGAAAT	TCGCCGTGGC	TTTCGTGCCT	CGAACGCAAA	GAACGACCCC	GCCAAGGAGA	CAGAGACCTT	CGTTGCCCAA	[990]
Ploeotia_oblonga_gDNA_2	ACGTTGGCTT	CAACGTCAAG	AACGTGGCCG	TCAAGGATCT	CCGCCGTGGG	TTTCGTGTCT	CCAACGACAA	GCAGGACCCC	GCCAAGGAGG	CTGAGGACTT	CACCGCTCAG	[990]
Peranema_trichophorum_FJ807243_2	AcGtTgGaTt	CAACGTAAAg	AaCgTcTcGg	TcAaGgATaT	CaGaAgTgGa	TtCgTtGcCt	CaAaCtCaAa	GaACGaTcCc	GcAaAgGaAa	CcGcTgAcTt	CaCtGcTcAa	[984]
Peranema_trichophorum_2	ACGTTGGATT	CAACGTAAAG	AACGTCTCCG	TCAAGGATAT	CAGACGTGGA	TTTCGTTGC	-----	-----	-----	-----	-----	[938]
Euglena_stellata_2	ATGTTGGCTT	CAACGTGAAG	AATGTGTCTG	TAAAGATAT	TCGACGAGGT	TTTCGTCTGT	CCAATGCCAA	GAATGACCCA	GCCAAGGAAG	CTGGTGATTT	CACGTGCTCAG	[870]
Euglena_hiemalis_1	RTGTTGGTTT	CAACGTCAAG	AACGTGTCTG	TCAAGGATAT	CCGTCGTGGC	TACGTCGCTT	CCAACGCAAA	GAACGATCCC	GCCAAGGAAG	CAGCAGATTT	CACAGCTCAG	[990]
Euglena_hiemalis_2	ACGTTGGCTT	CAACGTGAAG	AACGTGTCTG	TCAAGGATAT	CCGTCGTGGT	TATGTGGCCT	CCAACGCCAA	GAACGACCCC	GCCAAGGAGG	CAGCTGACTT	CACGTGCCAG	[990]
Euglena_exilis	ATGTTGGATT	CAACGTCAAG	AACGTATCTG	TCAAGGATAT	CCGTCGTGGA	TTTGTCTGTT	CCAATGCCAA	GAACGACCTT	GCCAAGGAGG	CATCAGATTT	CACGTGCACAG	[870]
Strombomonas_triquetra_1	ACGTTGGTTT	CAACRTGRAG	AACGTGTCTG	TTAAGGACAT	TCGCCGAGGG	TATGTCTGCT	CTAACGCAAA	GAATGATCCT	GCCAAGGAAG	CTGGAGATTT	CACGTGCGCAG	[990]
Strombomonas_napiformis_1	ACGTTGGTTT	CAACGTGAAG	AACGTGTCTG	TGAAGGACAT	TCGTCGAGGA	TATGTGTGCT	CAAATGCCAA	GAATGATCCC	GCCAAGGAGG	CTGCTGATTT	CACGTGCACAG	[990]
Strombomonas_acuminata_ACOI_2467	ACGTTGGTTT	CAATGTGAAG	AACGTATCCG	TGAAGGACAT	TCGTCGAGGA	TATGTGTGCT	CAAATGCCAA	GAATGATCCC	GCCAAGGAGG	CTGGTGATTT	CACGTGCACAG	[990]
Strombomonas_acuminata_SAG1280_1_1	ACGTTGGTTT	CAATGTGAAG	AACGTATCCG	TGAAGGACAT	TCGTCGAGGA	TATGTGTGCT	CAAATGCCAA	GAATGATCCC	GCCAAGGAGG	CTGGTGATTT	CACGTGCACAG	[872]
Phacus_orbicularis	ACGTTGGCTT	CAACGTCAAG	AACGTGTCTG	TCAAGGATAT	TCGTCGTGGA	TTTGTGGCCT	CCAATGCCAA	GAATGACCCA	GCCAAGGAGG	CTTCTGACTT	CACGTGCCAG	[948]
Euglena_stellata_2_2	ATGTTGGTTT	CAACGTCAAG	AACGTTTCTG	TGAAGGATAT	TCGCCGTGGA	TTTGTGTGTT	CCAACGCTAA	GAATGACCTT	GCCAAGGAAG	CTGGTGATTT	CACGTGCACAG	[867]
Strombomonas_similis	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TCAAGGATAT	TCGACGAGGT	TTTGTCTGCT	CCAATGCCAA	GAATGACCTT	GCCAAGGAAG	CTGGTGACTT	CACGTGCACAG	[870]
Trachelomonas_volvocina	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TCAAGGACAT	TCGTCGTGGG	TTTGTCTGCT	CAAACGCCAA	GAATGACCTT	GCCAAGGAGG	CTGGTGATTT	CACGTGCACAG	[990]
Euglena_clara	ACGTTGGTTT	CAACGTGAAG	AATGTGTCCG	TCAAGGATAT	TCGCCGTGGG	TTTCGTGCAAT	CTAATTCAAA	GAACGATCCC	GCCAAGGAGG	CTGCTGATTT	CACGTGCCAG	[870]
Euglena_deses	ACGTTGGTTT	CAATGTAAAG	AACGTGTCCG	TGAAGGATAT	CCGACGTGGC	TATGTTGCCCT	CCAATGCCAA	GAATGATCCT	GCCAAGGAGG	CGGGTGATTT	CACGTGCTCAG	[990]
Lepocinclis_ovum	ATGTTGGCTT	CAACGTGAAG	RACGTGTCTG	TCAAGGATAT	CCGCCGTGGC	TATGTCGTCT	CCAACCTCAA	GAATGACCTT	GCCAAGGAGG	CGGCTGATTT	CACGTGCTCAG	[870]
Euglena_granulata	ACGTTGGTTT	CAACGTGAAG	AACGTGTCTG	TGAAGGATAT	TCGACGTGGA	TATGTCGTCT	CTAACGCGAA	GAATGACCCC	GCCAAGGAAG	CTGCTGACTT	CACGTGCCAG	[990]
Trachelomonas_planctonica	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TCAAGGATAT	CCGACGTGGA	TATGTTGCTT	CCAATGCCAA	GAATGACCCC	GCCAAGGAGG	CTGCAGACTT	CACGTGCCAG	[872]
Trachelomonas_lefevrei	ATGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TGAAGGATAT	CCGACGTGGA	TATGTTGCCCT	CCAATGCCAA	GAACGACCTT	GCCAAGGAGG	CTGGGGACTT	CACGTGCCAG	[990]
Trachelomonas_echinata	ACGTTGGTTT	CAACGTGAAG	AACGTGTCTG	TCAAGGATAT	TCGACGAGGC	TATGTCGCGT	CGAATGCCAA	GAATGATCCA	GCCAAGGAGG	CTGGTGACTT	CACGTGCTCAG	[990]
Discoplastis_adunca	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TCAAGGATAT	TCGACGTGGC	TACGTGGCCT	CCAACGCCAA	GAATGACCTT	GCCAAGGAGG	CAGCTGATTT	CACGTGCTCAG	[990]
Lepocinclis_triapteris_1	ATGTTGGCTT	CAACGTGAAG	AACGTGTCTG	TCAAGGATAT	CCGTCGTGGT	TATGTTGCCCT	CCAATGCCAA	GAACGATCCC	GCCAAGGAAG	CTGGCGACTT	CACGTGCACAG	[990]
Lepocinclis_acus_var_major	ACGTTGGTTT	CAACGTGAAG	AACGTGGCCG	TCAAGGATAT	CCGCCGTGGC	TACGTTGCTT	CTAAGCTCAA	GAATGATCCG	GCCAAGGAGG	CTGCGGACTT	CACGTGCCAG	[990]
Lepocinclis_steinii	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TGAAGGATAT	CCGCCGTGGG	TTTGTCTGCT	CCAACGCTAA	GAATGATCCT	GCCAAGGAGG	CTGCTGATTT	CACAGCCCAAG	[870]
Monomorpha_aenigmatica_2	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TGAAGGATAT	CCGCCGTGGA	TACGTCGTCT	CCAATGCCAA	GAATGACCCA	GCCAAGGAGG	CTGCTGACTT	CACCGCCCAAG	[990]
Strombomonas_ovalis_2	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TGAAGGATAT	CCGCCGTGGG	TACGTCGTCT	CCAATGCCAA	GAATGACCCA	GCCAARGAGG	CTGCTGACTT	CACCGCCCAAG	[872]
Strombomonas_borysteniensis_3	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TGAAGGATAT	CCGCCGTGGA	TACGTCGTCT	CCAATGCCAA	GAATGACCCA	GCCAAGGAGG	CTGCTGACTT	CACCGCCCAAG	[862]
Phacus_pusillus	ACGTTGGCTT	CAACGTGAAG	AACGTGTCCG	TCAAGGATAT	CCGCCGTGGC	TACGTCGCTT	CCAACGCCAA	GAATGACCCC	GCCAAGGAGG	CTGGCGACTT	CACGCGCCAG	[990]
Lepocinclis_triapteris_2	ACGTTGGCTT	CGACGTCAAG	AACGTGTCCG	TCAAGGATAT	CCGCCGTGGC	TACGTCGCTT	CCAACGCCAA	GAATGACCCG	GCCAAGGAGG	CTGGCGACTT	CACGCGCCAG	[872]
Trachelomonas_conradi	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TCAAGGATAT	TCGACGTGGG	TATGTGGCCT	CCAACGCCAA	GAACGACCTT	GCCAAGGAAG	CTGGGGACTT	CACGTGCCAG	[990]
Monomorpha_sp	ATGTTGGCTT	CAACGTGAAG	AACGTTGCTG	TCAAGGACAT	CCGTCGTGGT	TACGTCGTGGT	CCAACGCCAA	GAATGACCTT	GCCAAGGAGG	CTGCAGACTT	TACCGCTCAG	[990]
Euglena_gracilis_gDNA_3	ATGTTGGCTT	CAACGTGAAG	AACGTTGCCG	TTAAGGATAT	TCGCCGTGGC	TACGTTGGCT	CCAACCTCAA	GAATGACCTT	GCCAAGGAGG	CAGCTGACTT	CACGTGCCAG	[990]
Eutreptiella_pomquetensis	ATGTTGGCTT	CAACGTGAAG	AACGTTGCCG	TTAAGGATAT	TCGCCGTGGC	TACGTTGGCT	CCAACCTCAA	GAATGACCTT	GCCAAGGAGG	CAGCTGACTT	CACGTGCCAG	[990]
Eutreptia_pertyi	ATGTTGGCTT	CAACGTGAAG	AACGTTGCCG	TTAAGGATAT	TCGCCGTGGC	TACGTTGGCT	CCAACCTCAA	GAATGACCTT	GCCAAGGAGG	CAGCTGACTT	CACGTGCCAG	[990]
Monomorpha_pyrum	ATGTTGGCTT	CAACGTGAAG	AACGTTGCCG	TTAAGGATAT	TCGCCGTGGC	TACGTTGGCT	CCAACCTCAA	GAATGACCTT	GCCAAGGAGG	CAGCTGACTT	CACGTGCCAG	[990]
Cryptoglena_pigra_gDNA_4	ATGTTGGCTT	CAACGTGAAG	AACGTTGCCG	TTAAGGATAT	TCGCCGTGGC	TACGTTGGCT	CCAACCTCAA	GAATGACCTT	GCCAAGGAGG	CAGCTGACTT	CACGTGCCAG	[990]
Euglena_gracilis_gDNA_2	ACGTTGGCTT	CAACGTGAAG	AACGTGTCTG	TCAAGGATAT	TCGTCGTGGT	TATGTGGCCT	CGAACGCCAA	GAACGATCCC	GCCAAGGAGG	CTGCTGACTT	CACGTGCTAG	[990]
Euglena_gracilis_1	ACGTTGGCTT	CAACGTGAAG	AACGTGTCTG	TCAAGGATAT	TCGTCGTGGT	TATGTGGCCT	CGAACGCCAA	GAACGATCCC	GCCAAGGAGG	CTGCTGACTT	CACGTGCCAG	[990]
Ploeotia_oblonga_1	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TCAAGGATAT	TCGTCGTGGG	TATGTGTGCT	CCAACGCCAA	GAACGACCTT	GCCAAGGAGG	CCGCCGACTT	CACCGCTCAG	[990]
Euglena_geniculata	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TCAAGGATAT	TCGTCGTGGG	TATGTGTGCT	CCAACGCCAA	GAACGACCTT	GCCAAGGAGG	CCGCCGACTT	CACCGCTCAG	[990]
Cryptoglena_pigra_2	ACGTTGGCTT	CAACGTGAAG	AACGTTGCTG	TCAAGGACAT	CCGTCGTGGG	TATGTCTGCT	CCAACCTCAA	GAATGATCCG	GCAAAGGAGG	CTGGTGATTT	CACGTGCTCAG	[990]
Cryptoglena_pigra_gDNA_3	ACGTTGGCTT	CAACGTGAAG	AACGTTGCTG	TCAAGGACAT	CCGCCGTGGC	TATGTCTGCT	CCAACCTCAA	GAACGATCCG	GCAAAGGAGG	CTGGTGACTT	CACGTGCTCAG	[990]
Cryptoglena_skujai_2	ACGTTGGCTT	CAACGTGAAG	AACGTTGCTG	TCAAGGACAT	CCGTCGTGGC	TATGTCTGCT	CCAACCTCAA	GAACGATCCG	GCAAAGGAGG	CTGGTGACTT	CACGTGCTCAG	[990]
Cryptoglena_pigra_1	ACGTTGGTTT	CAACGTGAAG	AACGTTGCTG	TGAAGGATAT	TCGTCGTGGG	TATGTGTGCT	CCAATTGCAA	GAACGATCCG	GCCAAGGAGG	CGGGTGACTT	CACGTGCCAG	[990]
Cryptoglena_skujai_1	ACGTTGGTTT	CAACGTGAAG	AACGTTGCTG	TGAAGGATAT	CCGACGTGGC	TTTGTGTGCT	CCAATTGCAA	GAACGATCCT	GCCAAGGAGG	CGGCTGACTT	CACGCGCCAG	[990]
Colacium_vesiculosum	ACGTTGGCTT	CAACGTGAAG	AACGTGTCCG	TGAAGGACAT	CCGACGTGGC	TATGTCTGCT	CCAACGCCAA	GAACGACCCC	GCAAAGGAGG	CCGCCGACTT	TACTGCCCAAG	[949]
Colacium_sp	ACGTTGGTTT	CAACGTGAAG	AACGTGTCCG	TCAAGGATAT	CCGTCGTGGG	TACGTTTGGT	GAACGACCCC	GAACGACCCC	GCGGAGGAGG	CCGGTGACTT	CACCGCCCAAG	[990]
Colacium_elongatum	ACGTTGGTTT	CAACGTGAAG	AACGTGTCTG	TCAAGGATAT	CCGACGTGGC	TACGTGGCTT	CTAATGCCAA	GAACGACCCC	GCAAAGGAGG	CTGGTGACTT	CACCGCCCAAG	[990]
Strombomonas_triquetra_2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[858]
Strombomonas_ovalis_1	ACGTCGGCTT	CAACGTGAAG	AACGTTTCCG	TCAAGGATAT	CCGCCGTGGC	TACGTCGTCT	CCAACGCCAA	GAACGATCCC	GCCAAGGAGG	CTGGTGACTT	CACGTGCTCAG	[990]
Monomorpha_aenigmatica_1	ACGTCGGCTT	CAACGTGAAG	AACGTTTCCG	TCAAGGATAT	CCGCCGTGGC	TACGTCGTCT	CCAACGCCAA	GAACGATCCC	GCCAAGGAGG	CTGGTGACTT	CACGTGCTCAG	[862]
Strombomonas_borysteniensis_1	ACGTCGGCTT	CAACGTGAAG	AACGTTTCCG	TCAAGGATAT	CCGCCGTGGC	TACGTCGTCT	CCAACGCCAA	GAACGATCCC	GCCAAGGAGG	CTGCTGACTT	CACGTGCTCAG	[870]
Strombomonas_borysteniensis_gDNA_2	ACGTCGGCTT	CAAC	-----	-----	-----	-----	-----	-----	-----	-----	-----	[894]
Strombomonas_napiformis_2	ATGTATGCTT	CAACGTGAAG	AACGTGTCTG	TGAAGGATAT	CCGTCGTGGA	TATGTGTGTT	CCAATGCCAA	GAACGATCCT	GCCAAGGAGG	CTGGTGACTT	CACCGCCCAAG	[990]
Strombomonas_napiformis_gDNA_3	ATGTAGGCTT	CAACGTGAAG	AACGTGTCTG	TGAAGGATAT	CCGTCGTGGA	TATGTGTGTT	CCAATGCCAA	GAACGATCCT	GCCAAGGAGG	CTGGTGACTT	CACCGCCCAAG	[990]
Strombomonas_acuminata_SAG1280_1_2	ACGTTGGCTT	CAACGTGAAG	AACGTGTCCG	TGAAGGATAT	CCGCCGAGGT	TATGTCTGCT	CCAATGCCAA	GAACGATCCG	GCCAAGGAGG	CTGGTGACTT	CACCGCTCAG	[872]

	1000	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	
Entosiphon_sulcatum_FJ807254	GTCATAATCA	TGAACCATCC	TGGACAGA--	-----	-----	-----	-----	-----	-----	-----	-----	[966]
Entosiphon_sp	GTCAATTGTC	TGAACCAACC	CGGTACAGATC	TCCAACGGCT	ACGCCCCGGT	GCTCGATTGC	CACACCTGCC	ACATTGCCTG	CAAGTTGCCT	GAGATCTTGA	CAAAGATCGA	[1100]
Ploeotia_oblonga_gDNA_2	GTCAATCATCT	TGAACCAACC	CGGTACAGATC	TCCAACGGGT	ACGCCCCCGT	CATCGATTGC	CACACGTGCC	ATATCGCGTG	CAAGTTGCCT	GAATCAAGA	CAAAGATTGA	[1100]
Peranema_trichophorum_FJ807243_2	GtCaTtGtTc	TtAaCcAcCc	TgTcCAAAcTc	CaAaAcGgAt	AcGcCcCaGt	GtTgGaTtGc	CaCaCtTcAc	AcAtTgCtTg	CaAaTtCtAc	GaGaTcAaGa	CcAaGaTcGa	[1094]
Peranema_trichophorum_2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[938]
Euglena_stellata_2	GTCAATCATTT	TGAACCATCC	TGGGCAAAAT	GGCAACGGTT	ATGCACCTGT	GTTGGACTGT	CACACATGCC	ACATTGCCTG	CRAGTTTGCC	ACAATTCCAGA	CCAAAATTGA	[980]
Euglena_hiemalis_1	GTCAATCGTCC	TGAACCAACC	AGGTCAAATT	GGAACCGGCT	ATGCGCCCGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[1100]
Euglena_hiemalis_2	GTTATCGTCC	TGAACCAACC	TGGCCAGATC	GGAACCGGCT	ATGCACCAAGT	CCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CAAAGATTGA	[1100]
Euglena_exilis	GTCAATCATTC	TGAACCAACC	AGGTACAGATT	GGCAACGGCT	ATGCTCCAAGT	GCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAAATTCGCC	ACCATCCAGA	CAAAGATTGA	[980]
Strombomonas_triquetra_1	GTCAATCATCC	TGAACCAACC	CGGGCAGATT	GGGAACGGGT	ACGCCCCCGT	CCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATCGA	[1100]
Strombomonas_napiformis_1	GTGATCATCC	TGARCCACCC	TGGTCAGATT	GGCAATGGCT	ATGCTCCTGT	CCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[1100]
Strombomonas_acuminata_ACOI_2467	GTGATCATCC	TGAACCAACC	TGGGCAGATT	GGCAACGGCT	ATGCCCTGT	CCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[1100]
Strombomonas_acuminata_SAG1280_1_1	GTGATCATCC	TGAACCAACC	TGGGCAGATT	GGCAACGGCT	ATGCCCTGT	CCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[982]
Phacus_orbicularis	GTCAATCGTCC	TGAATCATCC	TGGCCAGATT	GGGAACGGTT	ACGCACCAAGT	GCTGGATTGC	CACACTTGCC	ACATTGCATG	CAAGTTTGCA	GAGATTCCGA	CAAAGATTGA	[1058]
Euglena_stellata_2_2	GTTATCATCC	TGAATCACC	TGGCCAAAT	GGAATGGCT	ATGCACCTGT	GCTTGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAAA	CAAAGATTGA	[977]
Trachelomonas_similis	GTTATCATTC	TCAACCAACC	TGGCCAAAT	GGCAATGGAT	ATGCACCAAGT	ATTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[980]
Trachelomonas_volvocina	GTCAATCATTC	TGAATCACC	TGGTCAGATT	GGCAATGGAT	ATGCACCAAGT	TCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[1100]
Euglena_clara	GTCAATCATCC	TTAACCAACC	TGGCCAGATT	GGCAATGGCT	ACGCCCTGT	GCTCGACTGC	CACACTTGCC	ACATCGCTTG	CAAGTTTGCC	ACCATCCAAA	CCAAAATCGA	[980]
Euglena_deses	GTTATCATCC	TGAATCATCC	AGGCCAGATT	GGAATGGAT	ATGCTCCAAGT	ATTGGATTGC	CACACTTGCC	ACATCGCTTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATTGA	[1100]
Lepocinclis_ovum	GTTATCATCC	TTAACCAACC	TGGCCAGATT	GGGAATGGTT	ATGCCCAAGT	GCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCAATCCAGA	CAAAGATTGA	[980]
Euglena_granulata	GTTATCATTC	TGAATCACC	TGGTCAGATT	GGCAACGGCT	ATGCACCTGT	GCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATCGA	[1100]
Trachelomonas_planctonica	GTCAATCATCC	TCAACCATCC	AGGCCAGATT	GGGAATGGCT	ATGCCCCCGT	CCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCAATCCAGA	CAAAGATTGA	[982]
Trachelomonas_lefevrei	GTCAATCATCC	TGAACCAACC	AGGCCAAAT	GGGAACGGCT	ATGCCCTGT	CCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCA	ACCAATCCAGA	CAAAGATTGA	[1100]
Trachelomonas_echinata	GTTATCATCC	TGAACCAACC	TGGCCAGATT	GGAAATGGGT	ACGCACCTGT	GCTGGACTGC	CATACATGCC	ACATTGCCTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATCGA	[1100]
Discoplastis_adunca	GTCAATCATCC	TGAACCAACC	TGGCCAAAT	GGGAATGGCT	ACGCCCAAGT	GCTGGATTGC	CACACTTGCC	ACATTGCATG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[1100]
Lepocinclis_tripteris_1	GTGATTGTC	TGAACCAACC	TGGCCAAAT	GGAATGGTT	ATGCCCTGT	GCTGGATTGC	CATACCTGCC	ATATCGCTTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATTGA	[1100]
Lepocinclis_acus_var_major	GTCAATCATTC	TGAACCAACC	TGGCCAGATT	GGAAATGGTT	ATGCCCAAGT	TCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCAATCCAGA	CAAAGATTGA	[1100]
Lepocinclis_steinii	GTCAATCATTC	TGAACCAACC	TGGTCAGATT	GGCAACGGTT	ATGCCCTGT	GCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATTGA	[980]
Monomorphina_aenigmatica_2	GTTATCATTC	TGAACCAACC	TGGCCAGATT	GGAACCGGCT	ATGCCCCCGT	GCTGGATTGC	CACACTTGCC	ACATTGCATG	CAAGTTTGCC	ACTATCCAGA	CAAAGATTGA	[1100]
Strombomonas_ovalis_2	GTTATCATTC	TGAACCAACC	TGGCCAGATT	GGAACCGGCT	ATGCC---	-----	-----	-----	-----	-----	-----	[918]
Strombomonas_borysteniensis_3	GTTATCATTC	TGAACCAACC	TGGCCAGATT	GGAACCGGCT	ATGCCCC---	-----	-----	-----	-----	-----	-----	[909]
Phacus_pusillus	GTCAATCATCC	TCAACCAACC	CGGCCAGATC	GGGAACGGCT	ACGCCCCCGT	GCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CAAAGATTGA	[1100]
Lepocinclis_tripteris_2	GTCAATCATCC	TCAACCAACC	CGGCCAGATC	GGGAACGGCT	AC---	-----	-----	-----	-----	-----	-----	[914]
Trachelomonas_conradi	GTTATCATCC	TCAACCATCC	TGGCCAGATT	GGAAATGGCT	ATGCCCCCGT	GCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CAAAGATTGA	[1100]
Monomorphina_sp	GTTATCATCC	TTAACCATCC	TGGCCAGATT	GGCAATGGCT	ATGCCCCCGT	GCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CGAAAATTGA	[1100]
Euglena_gracilis_gDNA_3	GTTATTATCC	TCAACCAACC	TGGCCAGATT	GGCAACGGCT	ACGCCCTGT	GTTGGATTGC	CACACTTGCA	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CGAAGATTGA	[1100]
Eutreptiella_pomquetensis	GTTATTATCC	TCAACCAACC	TGGCCAGATT	GGCAACGGCT	ACGCCCTGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CGAAGATTGA	[1100]
Eutreptia_pertyi	GTTATTATCC	TCAACCAACC	TGGCCAGATT	GGCAACGGCT	ACGCCCTGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CGAAGATTGA	[1100]
Monomorphina_pyrum	GTTATTATCC	TCAACCAACC	TGGCCAGATT	GGCAACGGCT	ACGCCCTGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CGAAGATTGA	[1100]
Cryptoglena_pigra_gDNA_4	GTTATTATCC	TCAACCAACC	TGGCCAGATT	GGCAACGGCT	ACGCCCTGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CGAAGATTGA	[1100]
Euglena_gracilis_gDNA_2	GTTATCATTC	TGAACCAACC	TGGTCAGATT	GGCAACGGTT	ATGCCCCCGT	GCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CGAAGATTGA	[1100]
Euglena_gracilis_1	GTTATCATTC	TGAACCAACC	TGGTCAGATT	GGCAACGGTT	ATGCCCCCGT	GCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CGAAGATTGA	[1100]
Ploeotia_oblonga_1	GTCAATTATCC	TGAATCACC	TGGTCAGATC	GGAACCGGTT	ACGCACCAAGT	GCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAAA	ACAAGATTGA	[1100]
Euglena_geniculata	GTCAATTATCC	TGAATCACC	TGGTCAGATC	GGAACCGGTT	ACGCACCAAGT	GCTGGACTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAAA	ACAAGATTGA	[1100]
Cryptoglena_pigra_2	GTGATCATTC	TGAACCAACC	TGGCCAGATT	GGCAACGGCT	ATGCCCAAGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CAAAGATTGA	[1100]
Cryptoglena_pigra_gDNA_3	GTGATTATCC	TGAGCCACCC	AGGCCAGATT	GGCAATGGCT	ACGCCCAAGT	GCTGGATTGC	CACACTTGCC	ACATCGCTTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[1100]
Cryptoglena_skujai_2	GTGATTATCC	TGAACCAACC	AGGCCAGATT	GGCAATGGCT	ACGCCCAAGT	GCTGGATTGC	CACACTTGCC	ACATCGCTTG	CAAGTTTGCA	ACCATCCAGA	CAAAGATTGA	[1100]
Cryptoglena_pigra_1	GTCAATTATTC	TGAACCAACC	TGGCCAGATT	GGCAACGGCT	ACGCACCTGT	GCTGGATTGC	CACACTTGCC	ACATTGCTTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATCGA	[1100]
Cryptoglena_skujai_1	GTCAATCATTC	TGAACCAACC	TGGCCAGATT	GGCAATGGCT	ACGCACCCGT	GCTGGATTGC	CACACTTGCT	ACATTGCTTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATTGA	[1100]
Colacium_vesiculosum	GTCAATCATCC	TGAACCAACC	TGGTCAGATT	GGCAATGGCT	ATGCCCTGT	TTTGGATTGC	CACACTTGCC	ACATCGCTTG	TAAGTTTGCC	ACCATCCAGA	CAAAGATTGA	[1059]
Colacium_sp	GTCAATCATCC	TGAACCAACC	CGGCCAGATC	GGCAACGGTT	ATGCCCCCGT	GCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CAAAGATTGA	[1100]
Colacium_elongatum	GTCAATTATCC	TGAACCAACC	TGGTCAGATT	GGCAACGGTT	ATGCCCCCGT	TCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTGCCT	ACCATCCAGA	CAAAGATTGA	[1100]
Strombomonas_triquetra_2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[858]
Strombomonas_ovalis_1	GTCAATCATTC	TGAACCAACC	TGGTCAGATT	GGCAACGGCT	ATGCCCAAGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATTGA	[1100]
Monomorphina_aenigmatica_1	GTCAATCATTC	TGAACCAACC	TGGTCAGATT	GGCAACGGCT	ATGCCCAAGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATTGA	[972]
Strombomonas_borysteniensis_1	GTCAATCATTC	TGAACCAACC	TGGTCAGATT	GGCAACGGCT	ATGCCCAAGT	GTTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCATCCAGA	CAAAGATTGA	[980]
Strombomonas_borysteniensis_gDNA_2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[894]
Strombomonas_napiformis_2	GTTATCATCC	TCAACCAACC	TGGCCAGATT	GGAATGGCT	ATGCACCAAGT	GCTTGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCATCCAGA	CGAAGATTGA	[1100]
Strombomonas_napiformis_gDNA_3	GTTATCATCC	TCAACCAACC	TGGCCAGATT	GGAATGGCT	ATGCACCAAGT	GCTTGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCATCCAGA	CGAAGATTGA	[1100]
Strombomonas_acuminata_SAG1280_1_2	GTTATCATCC	TCAACCAACC	TGGCCAGATT	GGAATGGCT	ATGCACCAAGT	GCTGGATTGC	CACACTTGCC	ACATTGCCTG	CAAGTTTGCC	ACCATCCAGA	CGAAAATTGA	[982]

	1110	1120	1130	1140	1150	1160	
Entosiphon_sulcatum_FJ807254	-----	-----	-----	-----	-----	-----	[966]
Entosiphon_sp	CCGCCGCTCT	GGCAAGGAGC	TCGAGCAGGA	GCCCAAGGCT	ATCAAGTCCG	GTGATGCCGC	CATGGTC [1167]
Ploeotia_oblonga_gDNA_2	CCGTCGTTCC	GGCAAGGAGC	TCGAGGTTGA	GCCCAAGGCC	ATCAAGAGCG	GTGACGCCGC	TATCGTG [1167]
Peranema_trichophorum_FJ807243_2	CaGaCgTaGt	GgAaAgGaAt	TgGaAgCcGa	AcCaAaAtTc	AtCaAgAgTg	GaGaCgCtGc	CaTcGtC [1161]
Peranema_trichophorum_2	-----	-----	-----	-----	-----	-----	[938]
Euglena_stellata_2	TCGCCGTTCT	GGTAARGAGC	TCGAGGCTGA	GCCCAAGTTC	ATTAATCTG	GAGATGCTGC	TATTGTG [1047]
Euglena_hiemalis_1	CAGGCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCCG	GGGACGCCGC	CATTGTC [1167]
Euglena_hiemalis_2	CCGACGTTCT	GGGAAGGAGC	TGGAGGCAGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATTGTT [1167]
Euglena_exilis	TCGTGATCT	GGCAAGGAGC	TCGAGGCTGA	ACCCAAGTTC	ATCAAGTCAG	GAGATGCTGC	AATCGTG [1047]
Strombomonas_triquetra_1	CCGCCGCTCT	GGCAAGGAGT	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGACGCTGC	CATTGTG [1167]
Strombomonas_napiformis_1	CCGCCGTTCA	GGCAAGGAGC	TTGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GGGACGCTGC	CATCGTG [1167]
Strombomonas_acuminata_ACOI_2467	CCGCCGTTCA	GGCAAGGAGC	TTGAGGCCGA	ACCCAAGTTC	ATCAAGTCTG	GGGACGCTGC	CATTGTG [1167]
Strombomonas_acuminata_SAG1280_1_1	CCGCCGTTCA	GGCAAGGAGC	TTGAGGCCGA	ACCCAAGTTC	ATCAAGTCTG	GGGACGCTGC	CATTGTG [1049]
Phacus_orbicularis	TCGCCGTTCT	GGGAAGGAGT	TGGAGGCCGA	GCCCAAGTTC	ATCAATCTG	GAGATGCTGC	CATCGTT [1125]
Euglena_stellata_2_2	CCGTGATCT	GGCAAGGAGC	TGGAGGCAGA	GCCTAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATTGTT [1044]
Trachelomonas_similis	TCGTGTTCA	GGCAAGGART	TGGAGGCTGA	GCCCAAGTTC	ATCAATCTG	GAGATGCCGC	TATTGTT [1047]
Trachelomonas_volvocina	TCGCCGTTCT	GGCAAGGAGC	TCGAGGCCGA	GCCCAAGTTC	ATCAAGTCCG	GAGATGCCGC	CATTGTG [1167]
Euglena_clara	TCGACGATCT	GGCAAGGAGT	TGGAGGCTGA	GCCCAAGTTC	ATCAAGTCAG	GCGATGCTGC	TATCGTC [1047]
Euglena_deses	CCGACGTTCT	GGCAAGGAGT	TGGAGGCTGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCCGC	CATCGTT [1167]
Lepocinclis_ovum	CCGACGATCT	GGTAAGGAAT	TGGAGGCTGA	ACCCAAGTTC	ATCAAGTCTG	GCGATGCTGC	CATCGTG [1047]
Euglena_granulata	TCGCCGATCT	GGCAAGGAGC	TGGAGGCTGA	ACCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATCGTG [1167]
Trachelomonas_planctonica	CCGACGTTCT	GGCAAGGAGC	TGGAGGCTGA	GCCCAAGTTC	ATCAAGTCAG	GTGATGCCGC	CATTGTG [1049]
Trachelomonas_lefevrei	CCGACGTTCT	GGCAAGGAGC	TGGAGGCTGA	GCCCAAGTTC	ATCAAGTCAG	GTGATGCCGC	CATTGTG [1167]
Trachelomonas_echinata	CCGTGTTCT	GGCAAGGAGC	TGGAGGCAGA	ACCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATTGTG [1167]
Discoplastis_adunca	CCGACGTTCC	GGCAAGGAGC	TGGAGGCTGA	GCCCAATTC	ATCAAGTCTG	GTGATGCCGC	CATTGTC [1167]
Lepocinclis_tripteris_1	CCGACGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCAGC	CATCGTG [1167]
Lepocinclis_acus_var_major	CCGTGATCT	GGTAAGGAGC	TGGAGGCAGA	GCCCAAGCAC	ATCAAGTCTG	GTGATGCTGC	GATTGTG [1167]
Lepocinclis_steinii	TCGCCGTTCC	GGCAAGGAGC	TGGAGGCTGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATCGTG [1047]
Monomorpha_aenigmatica_2	CCGTGCTTCT	GGCAAGGAGT	TGGAGGCTGA	GCCCAAGTTC	ATCAATCTG	GTGATGCTGC	TATTGTC [1167]
Strombomonas_ovalis_2	-----	-----	-----	-----	-----	-----	[918]
Strombomonas_borysteniensis_3	-----	-----	-----	-----	-----	-----	[909]
Phacus_pusillus	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAATTC	ATCAAGTCTG	GTGATGCTGC	CATCGTG [1167]
Lepocinclis_tripteris_2	-----	-----	-----	-----	-----	-----	[914]
Trachelomonas_conradi	CCGACGTTCT	GGCAAGGAGC	TGGAGGCCGA	ACCTAAGTTC	ATCAAGTCCG	GTGATGCTGC	CATTGTG [1167]
Monomorpha_sp	CCGACGTTCT	GGCAAGGAGC	TCGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCCGC	CATTGTG [1167]
Euglena_gracilis_gDNA_3	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAGTTC	ATCAAGTCTG	GTGATGCCGC	CATTGTG [1167]
Eutreptiella_pomquetensis	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAGTTC	ATCAAGTCTG	GTGATGCCGC	CATTGTG [1167]
Eutreptia_pertyi	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAGTTC	ATCAAGTCTG	GTGATGCCGC	CATTGTG [1167]
Monomorpha_pyrum	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAGTTC	ATCAAGTCTG	GTGATGCCGC	CATTGTG [1167]
Cryptoglena_pigra_gDNA_4	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAGTTC	ATCAAGTCTG	GTGATGCCGC	CATTGTG [1167]
Euglena_gracilis_gDNA_2	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATTGTT [1167]
Euglena_gracilis_1	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATCGTT [1167]
Ploeotia_oblonga_1	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATCGTT [1167]
Euglena_geniculata	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATCGTT [1167]
Cryptoglena_pigra_2	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCAGA	GCCCAAGTTC	ATCAAGTCTG	GGGATGCTGC	CATTGTG [1167]
Cryptoglena_pigra_gDNA_3	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GGGATGCCGC	CATCGTG [1167]
Cryptoglena_skujai_2	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GGGATGCCGC	CATCGTG [1167]
Cryptoglena_pigra_1	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	AATTGTG [1167]
Cryptoglena_skujai_1	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GCGATGCCGC	CATCGTG [1167]
Colacium_vesiculosum	CCGTGCTCT	GGCAAGGAGT	TGGAGGCTGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATTGTT [1126]
Colacium_sp	CCGTGTTCT	GGTAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCCGC	CATCGTC [1167]
Colacium_elongatum	CCGACGATCT	GGAAAGGAGT	TGGAGGCTGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	AATTGTG [1167]
Strombomonas_triquetra_2	-----	-----	-----	-----	-----	-----	[858]
Strombomonas_ovalis_1	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAGTTC	ATCAAGTCTG	GTGATGCTGC	TATCGTT [1167]
Monomorpha_aenigmatica_1	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	TATCGTT [1039]
Strombomonas_borysteniensis_1	CCGCCGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCTGC	CATCGTT [1047]
Strombomonas_borysteniensis_gDNA_2	-----	-----	-----	-----	-----	-----	[894]
Strombomonas_napiformis_2	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAGTTC	ATCAAGTCTG	GTGACGCAGC	CATCGTG [1167]
Strombomonas_napiformis_gDNA_3	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCTAAGTTC	ATCAAGTCTG	GTGACGCAGC	CATCGTG [1167]
Strombomonas_acuminata_SAG1280_1_2	CCGTGTTCT	GGCAAGGAGC	TGGAGGCCGA	GCCCAAGTTC	ATCAAGTCTG	GTGATGCAGC	CATTGTG [1049]

Figure 4.2. The interleaved alignment of kinetoplastid EF1 α paralogs, as determined by OrthoMCL on TriTrypDB v1.1. The *Trypanosoma* species and paralogs are listed first, followed by those of *Leishmania*. Only the first member of the genus has its complete name listed, whereas the others that follow it are abbreviated to just the capitalized first letter of the genus name followed directly by the specific epithet and the identifying sequence details from TriTrypDB. The column positions, from 1-1352, are given in ten base increments at the top of each page. The number of residues from the original sequence is given in brackets to the right of each line corresponding to its respective OTU.

	10	20	30	40	50	60	70	80	90	100	
Trypanosoma brucei_10_70_5650	ATGGGAAAGG	AAAAGGTGCA	CATGAATCTT	GTGGTGGTGG	GCCACGTCGA	TGCCGGTAAA	TCCACTGCAA	CGGGTCACTT	GATCTACAAG	TGCGGTGGTA	[100]
Tbrucei_10_70_5670	ATGGGAAAGG	AAAAGGTGCA	CATGAATCTT	GTGGTGGTGG	GCCACGTCGA	TGCCGGTAAA	TCCACTGCAA	CGGGTCACTT	GATCTACAAG	TGCGGTGGTA	[100]
Tbrucei_10_70_5680	-----	-----	-ATGAACCTC	GTGGTGGTGG	GCCACGTCGA	TGCCGGGCAAG	TCCACAGCCA	CTGGCCACCT	CATCTACAAG	TGCGGAGGCA	[0]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	ATGGGGAAGG	AAAAGGTGCA	CATGAACCTT	GTGGTGGTGG	GCCACGTCGA	TGCCGGGCAAG	TCCACAGCCA	CTGGCCACCT	CATCTACAAG	TGCGGAGGCA	[79]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Leishmania major_F17_0080	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Lmajor_F17_0081	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Lmajor_F17_0082	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Lmajor_F17_0083	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Lmajor_F17_0084	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Lmajor_F17_0085	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Lmajor_F17_0086	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Infantum_J17_V3_0090	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Infantum_J17_V3_0100	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Infantum_J17_V3_0110	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Infantum_J17_V3_0170	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Infantum_J17_V3_0180	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Infantum_J17_V3_0190	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
Infantum_J17_V3_0200	ATGGGCAAGG	ATAAGGTGCA	CATGAACCTT	GTGGTCTGTCG	GCCATGTGCA	CGCCGGGCAAG	TCCACCGCCA	CTGGCCACTT	GATCTACAAG	TGCGGTGGCA	[100]
	110	120	130	140	150	160	170	180	190	200	
Trypanosoma brucei_10_70_5650	TTGACAAACG	TACGATCGAG	AAGTTCGAGA	AAGAAAGCTG	CGACATTGGT	AAGGCGTCTAT	TCAAGTACGC	ATGGGTGCTGT	GACAAGCTGA	AGGCTGAGCG	[200]
Tbrucei_10_70_5670	TTGACAAACG	TACGATCGAG	AAGTTCGAGA	AAGAAAGCTG	CGACATTGGT	AAGGCGTCTAT	TCAAGTACGC	ATGGGTGCTGT	GACAAGCTGA	AGGCTGAGCG	[200]
Tbrucei_10_70_5680	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	TTGACAAGCG	GACGATTGAG	AAGTTCGAGA	AGGAGGCGGC	TGAAATCGCG	AAGTCCTCGT	TCAAGTACGC	GTGGGTGTTG	GACAAGCTGA	AGGCTGAGCG	[179]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	TTGACAAGCG	GACGATTGAG	AAGTTCGAGA	AGGAGGCGGC	TGAAATCGCG	AAGTCCTCGT	TCAAGTACGC	GTGGGTGTTG	GACAAGCTGA	AGGCTGAGCG	[200]
Leishmania major_F17_0080	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Lmajor_F17_0081	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Lmajor_F17_0082	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Lmajor_F17_0083	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Lmajor_F17_0084	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Lmajor_F17_0085	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Lmajor_F17_0086	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Infantum_J17_V3_0090	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Infantum_J17_V3_0100	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Infantum_J17_V3_0110	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Infantum_J17_V3_0170	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Infantum_J17_V3_0180	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Infantum_J17_V3_0190	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
Infantum_J17_V3_0200	TCGACAAGCG	CACGATCGAG	AAGTTCGAGA	AGGAGGCCGC	CGAGATCGGC	AAGGCGTCTT	TCAAGTACGC	GTGGGTGCTC	GACAAGCTGA	AGGCGGAGCG	[200]
	210	220	230	240	250	260	270	280	290	300	
Trypanosoma brucei_10_70_5650	CGAACGTGGT	ATCAGCATCG	ACATTGCAC	GTGGAAATTC	GAGTCACCCA	AGTCTGTCTT	CACTATTATT	GATGCTCCTG	GGCACCGTGA	CTTCATCAAG	[300]
Tbrucei_10_70_5670	CGAACGTGGT	ATCAGCATCG	ACATTGCAC	GTGGAAATTC	GAGTCACCCA	AGTCTGTCTT	CACTATTATT	GATGCTCCTG	GGCACCGTGA	CTTCATCAAG	[300]
Tbrucei_10_70_5680	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[0]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GACGCCCCCG	GGCACCGCGA	CTTCATCAAG	[279]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Leishmania major_F17_0080	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Lmajor_F17_0081	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Lmajor_F17_0082	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Lmajor_F17_0083	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Lmajor_F17_0084	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Lmajor_F17_0085	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Lmajor_F17_0086	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Infantum_J17_V3_0090	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Infantum_J17_V3_0100	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Infantum_J17_V3_0110	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Infantum_J17_V3_0170	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Infantum_J17_V3_0180	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Infantum_J17_V3_0190	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]
Infantum_J17_V3_0200	CGAGCGCGGC	ATCAGCATCG	ACATTGCGCT	GTGGAAAGTTT	GAGTCGCCCA	AGTCTGTGTT	CACGATCATC	GATGCCCCCG	GGCACCGCGA	CTTCATCAAG	[300]

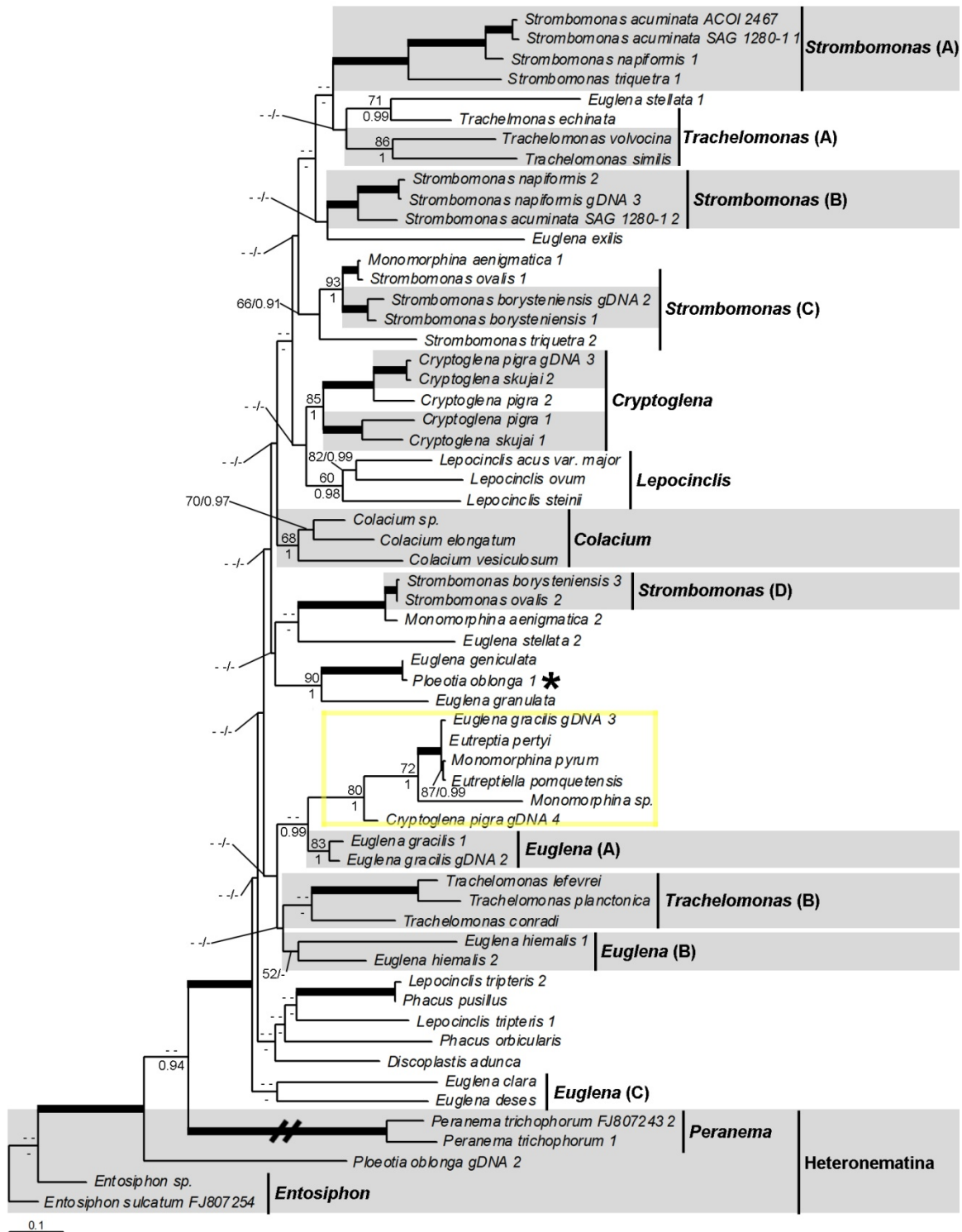
	310	320	330	340	350	360	370	380	390	400	
Trypanosoma_brucei_10_70_5650	AACATGATCA	CCGGCACATC	GCAAGCCGAC	GCAGCCATCC	TCATCATTTGC	CTCTGCGCAG	GGTGAGTTTC	AGGCTGGTAT	CTCCAAGGAT	GGACAGACCC	[400]
Tbrucei_10_70_5670	AACATGATCA	CCGGCACATC	GCAAGCCGAC	GCAGCCATCC	TCATCATTTGC	CTCTGCGCAG	GGTGAGTTTC	AGGCTGGTAT	CTCCAAGGAT	GGACAGACCC	[400]
Tbrucei_10_70_5680	--ATGATCA	CCGGCACATC	GCAAGCCGAC	CATCCATTTGC	TCATCATTTGC	CTCTGCGCAG	GGTGAGTTTC	AGGCTGGTAT	CTCCAAGGAT	GGACAGACCC	[97]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	AACATGATCA	CGGGCAGCTC	TCAGGCGGAC	CGCCGCGTCC	TGTTCATTGC	GTCAATCGCA	GGTGAGTTTC	AGGCGGGCAT	CTCGAAGGAC	GGCCAGACAC	[379]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	AACATGATCA	CGGGCAGCTC	TCAGGCGGAC	CGCCGCGTCC	TGTTCATTGC	GTCAATCGCA	GGTGAGTTTC	AGGCGGGCAT	CTCGAAGGAC	GGCCAGACAC	[400]
Leishmania_major_F17_0080	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Lmajor_F17_0081	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Lmajor_F17_0082	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Lmajor_F17_0083	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Lmajor_F17_0084	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Lmajor_F17_0085	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Lmajor_F17_0086	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Linfantum_J17_V3_0090	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Linfantum_J17_V3_0100	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Linfantum_J17_V3_0110	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Linfantum_J17_V3_0170	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Linfantum_J17_V3_0180	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Linfantum_J17_V3_0190	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
Linfantum_J17_V3_0200	AACATGATCA	CGGGCAGCTC	GCAGGCGGAC	CGCCGCAATC	TGATGATCGA	CTCGACGCAT	GGCGGCTTCG	AGGCTGGCAT	CTCGAAGGAC	GGCCAGACCC	[400]
	410	420	430	440	450	460	470	480	490	500	
Trypanosoma_brucei_10_70_5650	CGGAGCAGCG	GTTCGCTGGCC	TTCACCTTGG	GTGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGACTGTG	AACTACGGAC	AGGAGCGGTA	[500]
Tbrucei_10_70_5670	CGGAGCAGCG	GTTCGCTGGCC	TTCACCTTGG	GTGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGACTGTG	AACTACGGAC	AGGAGCGGTA	[500]
Tbrucei_10_70_5680	CGGAGCAGCG	GTTCGCTGGCC	TTCACCTTGG	GTGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGACTGTG	AACTACGGAC	AGGAGCGGTA	[500]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	CGGAGCAGCG	GTCTCTGGCC	TTCACGCTCG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	AACCTTCGCC	AGGAGCGGTA	[479]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	CGGAGCAGCG	GTCTCTGGCC	TTCACGCTCG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	AACCTTCGCC	AGGAGCGGTA	[500]
Leishmania_major_F17_0080	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Lmajor_F17_0081	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Lmajor_F17_0082	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Lmajor_F17_0083	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Lmajor_F17_0084	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Lmajor_F17_0085	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Lmajor_F17_0086	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Linfantum_J17_V3_0090	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Linfantum_J17_V3_0110	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Linfantum_J17_V3_0170	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Linfantum_J17_V3_0180	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Linfantum_J17_V3_0190	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
Linfantum_J17_V3_0200	CGGAGCAGCG	GTCTCTGGCC	TTCACCTTGG	GCGTGAAGCA	GATGGTTGTG	TGCTGCAACA	AGATGGACGA	CAAGTCCGTTG	ACGTACCGCG	AGTACGCGTA	[500]
	510	520	530	540	550	560	570	580	590	600	
Trypanosoma_brucei_10_70_5650	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[600]
Tbrucei_10_70_5670	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[600]
Tbrucei_10_70_5680	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[297]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	CGATGAGATT	GTGAAGGAGG	TGTCGGCGTA	CCTGAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	ATCCCATGCT	CGGGCTGGCA	GGGCGACAAC	[579]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	CGATGAGATT	GTGAAGGAGG	TGTCGGCGTA	CCTGAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	ATCCCATGCT	CGGGCTGGCA	GGGCGACAAC	[600]
Leishmania_major_F17_0080	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0081	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0082	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0083	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0084	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0085	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0086	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0090	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0100	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0110	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0170	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0180	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0190	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0200	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
	510	520	530	540	550	560	570	580	590	600	
Trypanosoma_brucei_10_70_5650	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[600]
Tbrucei_10_70_5670	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[600]
Tbrucei_10_70_5680	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[297]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	CGATGAGATT	GTGAAGGAGG	TGTCGGCGTA	CCTGAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	ATCCCATGCT	CGGGCTGGCA	GGGCGACAAC	[579]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	CGATGAGATT	GTGAAGGAGG	TGTCGGCGTA	CCTGAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	ATCCCATGCT	CGGGCTGGCA	GGGCGACAAC	[600]
Leishmania_major_F17_0080	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0081	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0082	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0083	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0084	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0085	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0086	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0090	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0100	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0110	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0170	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0180	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0190	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0200	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
	510	520	530	540	550	560	570	580	590	600	
Trypanosoma_brucei_10_70_5650	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[600]
Tbrucei_10_70_5670	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[600]
Tbrucei_10_70_5680	TGACGAGATT	GTGAAGGAGG	TGTCGTGCTTA	CATCAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	[297]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	CGATGAGATT	GTGAAGGAGG	TGTCGGCGTA	CCTGAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	ATCCCATGCT	CGGGCTGGCA	GGGCGACAAC	[579]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	CGATGAGATT	GTGAAGGAGG	TGTCGGCGTA	CCTGAAGAAG	GTTGGGTACA	ACGTGGAGAA	GGTGCGCTTC	ATCCCATGCT	CGGGCTGGCA	GGGCGACAAC	[600]
Leishmania_major_F17_0080	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0081	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0082	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0083	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0084	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0085	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Lmajor_F17_0086	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0090	CGATGAGATC	AGCAAAGGAG	TGGGCGCGTA	CCTGAAGCGC	GTTGGGTACA	ACCCGGAGAA	GGTGCGCTTC	ATCCCGATCT	CGGGCTGGCA	GGGCGACAAC	[600]
Linfantum_J17_V3_0100	CGATGAGATC	AGCAAAGGAG	TGGG								

	610	620	630	640	650	660	670	680	690	700	
Trypanosoma_brucei_10_70_5650	ATGATTGAGA	A-ATCCGAGA	AGATGCCATG	GTACAAGGGT	CCAACGCTCC	TGGAGGCACT	AGACATGCTG	GAG-CCACCA	GTGCGTCCGA	GCGACAAGCC	[698]
Tbrucei_10_70_5670	ATGATTGAGA	A-ATCCGAGA	AGATGCCATG	GTACAAGGGT	CCAACGCTCC	TGGAGGCACT	AGACATGCTG	GAG-CCACCA	GTGCGTCCGA	GCGACAAGCC	[698]
Tbrucei_10_70_5680	ATGATTGAGA	A-ATCCGAGA	AGATGCCATG	GTACAAGGGT	CCAACGCTCC	TGGAGGCACT	AGACATGCTG	GAG-CCACCA	GTGCGTCCGA	GCGACAAGCC	[395]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	ATGATTGACA	A-GTCGGAAA	ATATGCCGTG	GTACAAGGGC	CCCACGCTGC	TGGAGGCACT	CGACATGCTG	GAG-CCCCCG	GTGCGCCCCA	GCGACAAGCC	[677]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	ATGATTGAGA	A-GTCGGAAA	ATATGCCGTG	GTACAAGGGC	CCCACGCTGC	TGGAGGCACT	CGACATGCTG	GAG-CCCCCG	GTGCGCCCCA	GCGACAAGCC	[698]
Leishmainia_major_F17_0080	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Lmajor_F17_0081	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Lmajor_F17_0082	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	C6GATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Lmajor_F17_0083	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Lmajor_F17_0084	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Lmajor_F17_0085	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Lmajor_F17_0086	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Linfantum_J17_V3_0090	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Linfantum_J17_V3_0100	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Linfantum_J17_V3_0110	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Linfantum_J17_V3_0170	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Linfantum_J17_V3_0180	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Linfantum_J17_V3_0190	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
Linfantum_J17_V3_0200	ATGATCGAGA	A-GTCGGACA	ACATGCCGTG	GTACAAGGGT	CCCACGCTGC	TGGACGCGCT	CGACATGCTG	GAG-CCGCCG	GTGCGCCCGG	TGGACAAGCC	[698]
	710	720	730	740	750	760	770	780	790	800	
Trypanosoma_brucei_10_70_5650	CCTGCGTCTG	CCACTGCAAG	ACGTGTACAA	GATCGGTGGT	ATTGGCACCG	TGCCCGTTGG	TCGTGTGGAG	ACCGGGCTGA	TGAAGCCTGG	TGATGTGGTG	[798]
Tbrucei_10_70_5670	CCTGCGTCTG	CCACTGCAAG	ACGTGTACAA	GATCGGTGGT	ATTGGCACCG	TGCCCGTTGG	TCGTGTGGAG	ACCGGGCTGA	TGAAGCCTGG	TGATGTGGTG	[798]
Tbrucei_10_70_5680	CCTGCGTCTG	CCACTGCAAG	ACGTGTACAA	GATCGGTGGT	ATTGGCACCG	TGCCCGTTGG	TCGTGTGGAG	ACCGGGCTGA	TGAAGCCTGG	TGATGTGGTG	[495]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	GCTGCGCTCG	CCGCTGCAAG	ACGTGTACAA	GATCGGCTGG	ATCGGGACCG	TGCCGGTGGG	TCGCTGGAG	ACGGGACAG	TGAAGCCCGG	GACGCTGGTG	[777]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	GCTGCGCTCG	CCGCTGCAAG	ACGTGTACAA	GATCGGCTGG	ATCGGGACCG	TGCCGGTGGG	TCGCTGGAG	ACGGGACAG	TGAAGCCCGG	GACGCTGGTG	[798]
Leishmainia_major_F17_0080	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACGGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Lmajor_F17_0081	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Lmajor_F17_0082	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Lmajor_F17_0083	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Lmajor_F17_0084	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Lmajor_F17_0085	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Lmajor_F17_0086	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Linfantum_J17_V3_0090	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Linfantum_J17_V3_0110	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Linfantum_J17_V3_0170	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Linfantum_J17_V3_0180	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Linfantum_J17_V3_0190	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
Linfantum_J17_V3_0200	GCTGCGCTCG	CCCTGCAAG	ACGTGTACAA	GATCGGCGGT	ATCGGGACCG	TGCCCGTGGG	CCGCTGGAG	ACCGGATCA	TGAAGCCCGG	GACGCTGGTG	[798]
	810	820	830	840	850	860	870	880	890	900	
Trypanosoma_brucei_10_70_5650	ACGTTTGGCC	CGGCCAACGT	GACGACCAG	GTGAAATCGA	TCGAGATGCA	CCACGAGCAG	CTCGCTGAGG	GCACCCCCGG	TGACAAACGTC	GGCTTTAAGC	[898]
Tbrucei_10_70_5670	ACGTTTGGCC	CGGCCAACGT	GACGACCAG	GTGAAATCGA	TCGAGATGCA	CCACGAGCAG	CTCGCTGAGG	GCACCCCCGG	TGACAAACGTC	GGCTTTAAGC	[898]
Tbrucei_10_70_5680	ACGTTTGGCC	CGGCCAACGT	GACGACCAG	GTGAAATCGA	TCGAGATGCA	CCACGAGCAG	CTCGCTGAGG	GCACCCCCGG	TGACAAACGTC	GGCTTTAAGC	[595]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_10	ACGTTTGGCG	CGGCCAACGT	GACGACGAG	GTGAAAGTCGA	TTGAGATGCA	CCACGAGCAG	CTGGTTCGAGG	CCACGCCCGG	GACAAACGTC	GGCTTTCAAGC	[877]
Tcruzi_CLBrenner_Esmeraldo_like_1047053511369_20	ACGTTTGGCG	CGGCCAACGT	GACGACGAG	GTGAAAGTCGA	TTGAGATGCA	CCACGAGCAG	CTGGTTCGAGG	CCACGCCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Leishmainia_major_F17_0080	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Lmajor_F17_0081	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Lmajor_F17_0082	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Lmajor_F17_0083	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Lmajor_F17_0084	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Lmajor_F17_0085	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Lmajor_F17_0086	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Linfantum_J17_V3_0090	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Linfantum_J17_V3_0100	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Linfantum_J17_V3_0110	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Linfantum_J17_V3_0170	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Linfantum_J17_V3_0180	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Linfantum_J17_V3_0190	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]
Linfantum_J17_V3_0200	ACGTTTGGCG	CGGCCAACGT	GACGACTGAG	GTGAAAGTCGA	TCGAGATGCA	CCACGAGCAG	CTGGCGGAGG	GCAGAGCCGG	GACAAACGTC	GGCTTTCAAGC	[898]

	910	920	930	940	950	960	970	980	990	1000	
Trypanosoma_brucei_10_70_5650	TGAAGAACGT	TTCTGTAAAG	GACATCCGCC	GTGGCAACGT	CTGCGGTAAAC	ACCAAGAAGC	ACCCCCCAA	GGAGGCCGCC	GACTTCACGG	CACAGGTGAT	[998]
Tbrucei_10_70_5670	TGAAGAACGT	TTCTGTAAAG	GACATCCGCC	GTGGCAACGT	CTGCGGTAAAC	ACCAAGAAGC	ACCCCCCAA	GGAGGCCGCC	GACTTCACGG	CACAGGTGAT	[998]
Tbrucei_10_70_5680	TGAAGAACGT	TTCTGTAAAG	GACATCCGCC	GTGGCAACGT	CTGCGGTAAAC	ACCAAGAAGC	ACCCCCCAA	GGAGGCCGCC	GACTTCACGG	CACAGGTGAT	[695]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_10	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCCCCAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[977]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_20	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCCCCAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Leishmania_major_F17_0080	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGTAACTG	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Lmajor_F17_0081	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGTAACTG	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Lmajor_F17_0082	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGTAACTG	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Lmajor_F17_0083	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGTAACTG	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Lmajor_F17_0084	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGTAACTG	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Lmajor_F17_0085	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGTAACTG	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Lmajor_F17_0086	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGTAACTG	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Linfantum_j17_V3_0090	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Linfantum_j17_V3_0100	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Linfantum_j17_V3_0110	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Linfantum_j17_V3_0170	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Linfantum_j17_V3_0180	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Linfantum_j17_V3_0190	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
Linfantum_j17_V3_0200	TGAAGAACGT	GTCGGTGAAG	GACATCCGCC	GTGGCAACGT	GTGCGGCAAC	TCGAAGAAGC	ACCCGCCGAA	GGAAGCGGCC	GACTTCACGG	CGCAGGTGAT	[998]
	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	
Trypanosoma_brucei_10_70_5650	CATCCTGAAC	CACCCCGGAC	AGATTGAA	CGGTTATGCG	CCCGTGCTGG	ACTGCCACAC	ATCGCACATT	GCCTGCAAGT	TCGCGGAGAT	CGAGTCSAAG	[1098]
Tbrucei_10_70_5670	CATCCTGAAC	CACCCCGGAC	AGATTGAA	CGGTTATGCG	CCCGTGCTGG	ACTGCCACAC	ATCGCACATT	GCCTGCAAGT	TCGCGGAGAT	CGAGTCSAAG	[1098]
Tbrucei_10_70_5680	CATCCTGAAC	CACCCCGGAC	AGATTGAA	CGGTTATGCG	CCCGTGCTGG	ACTGCCACAC	ATCGCACATT	GCCTGCAAGT	TCGCGGAGAT	CGAGTCSAAG	[795]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_10	CATCCTGAAC	CACCCCGGCC	AGATCGCAAC	CGGCTATGCG	CCGGTGCTCG	ACTGCCACAC	CTGCCACATC	CGCTGCAAGT	TCGCCGAGAT	CGAGTCCAAAG	[1077]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_20	CATCCTGAAC	CACCCCGGCC	AGATCGCAAC	CGGCTATGCG	CCGGTGCTCG	ACTGCCACAC	CTGCCACATC	CGCTGCAAGT	TCGCCGAGAT	CGAGTCCAAAG	[1098]
Leishmania_major_F17_0080	CGTGCTGAAC	CACCCGGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Lmajor_F17_0081	CGTGCTGAAC	CACCCGGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Lmajor_F17_0082	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Lmajor_F17_0083	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Lmajor_F17_0084	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Lmajor_F17_0085	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Lmajor_F17_0086	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Linfantum_j17_V3_0090	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Linfantum_j17_V3_0100	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Linfantum_j17_V3_0110	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Linfantum_j17_V3_0170	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Linfantum_j17_V3_0180	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Linfantum_j17_V3_0190	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
Linfantum_j17_V3_0200	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	CGGCTATGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATT	TCGCCGAAAT	CGAGTCCAAAG	[1098]	
	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	
Trypanosoma_brucei_10_70_5650	ATCGACCGCTC	GCTCTGGCAA	GGAGCTGGAG	AAGGCTCCCA	AGTCGATCAA	GTCTGGCGAC	GCCGCGATCG	TGCGCATGGT	GCCGCGAAG	CCTATGTGCG	[1198]
Tbrucei_10_70_5670	ATCGACCGCTC	GCTCTGGCAA	GGAGCTGGAG	AAGGCTCCCA	AGTCGATCAA	GTCTGGCGAC	GCCGCGATCG	TGCGCATGGT	GCCGCGAAG	CCTATGTGCG	[1198]
Tbrucei_10_70_5680	ATCGACCGCTC	GCTCTGGCAA	GGAGCTGGAG	AAGGCTCCCA	AGTCGATCAA	GTCTGGCGAC	GCCGCGATCG	TGCGCATGGT	GCCGCGAAG	CCTATGTGCG	[895]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_10	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGCGCATGGT	GCCGCGAAG	CCGATGTGCG	[1177]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_20	ATCGACCGCC	GCTCGGGCAA	GGAGCTGGAG	AAGAACCCTA	AGTCGATCAA	GTCTGGGTAG	GCCGCGATCG	TGCGCATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Leishmania_major_F17_0080	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Lmajor_F17_0081	ATCGACCGCC	GCTCGGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Lmajor_F17_0082	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Lmajor_F17_0083	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Lmajor_F17_0084	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Lmajor_F17_0085	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Lmajor_F17_0086	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Linfantum_j17_V3_0090	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Linfantum_j17_V3_0100	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Linfantum_j17_V3_0110	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Linfantum_j17_V3_0170	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Linfantum_j17_V3_0180	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Linfantum_j17_V3_0190	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]
Linfantum_j17_V3_0200	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	AAGAACCCTA	AGGCGATCAA	GTCTGGCGAT	GCCGCGATCG	TGAAGATGGT	GCCGCGAAG	CCGATGTGCG	[1198]

	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	
Trypanosoma_brucei_10_70_5650	TGGAGGTCTT	CAACGACTAC	GCGCCACTCG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CCGTGCTGT	CGGTATCATC	AAGGCCGTGA	CCAAGAAGGA	[1298]
Tbrucei_10_70_5670	TGGAGGTCTT	CAACGACTAC	GCGCCACTCG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CCGTGCTGT	CGGTATCATC	AAGGCCGTGA	CCAAGAAGGA	[1298]
Tbrucei_10_70_5680	TGGAGGTCTT	CAACGACTAC	GCGCCACTCG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CCGTGCTGT	CGGTATCATC	AAGGCCGTGA	CCAAGAAGGA	[995]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_10	TGGAGGTGTT	CAACGACTAC	GCTCCTCTTG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	CGGCATCATC	AAGGCCGTGA	CGAAGAAGGA	[1277]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_20	TGGAGGTGTT	CAACGACTAC	GCTCCTCTTG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	CGGCATCATC	AAGGCCGTGA	CGAAGAAGGA	[1298]
Leishmania_major_F17_0080	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Lmajor_F17_0081	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Lmajor_F17_0082	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Lmajor_F17_0083	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Lmajor_F17_0084	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Lmajor_F17_0085	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Lmajor_F17_0086	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Linfantum_J17_V3_0090	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Linfantum_J17_V3_0100	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Linfantum_J17_V3_0110	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Linfantum_J17_V3_0170	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Linfantum_J17_V3_0180	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Linfantum_J17_V3_0190	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
Linfantum_J17_V3_0200	TGGAGGTGTT	CAACGACTAC	GCGCCGCTGG	GCCGCTTTGC	CGTGCCTGAC	ATGCGCCAGA	CGGTGCTGT	GGGCATCATC	AAGGCCGTGA	ACAAGAAGGA	[1298]
	1310	1320	1330	1340	1350						
Trypanosoma_brucei_10_70_5650	CGGTTCTGGT	GGTAAGGTGA	CGAAGGCTGC	GGTGAAGGCT	TCGAAGAAAT	AA					[1350]
Tbrucei_10_70_5670	CGGTTCTGGT	GGTAAGGTGA	CGAAGGCTGC	GGTGAAGGCT	TCGAAGAAAT	AA					[1350]
Tbrucei_10_70_5680	CGGTTCTGGT	GGTAAGGTGA	CGAAGGCTGC	GGTGAAGGCT	TCGAAGAAAT	AA					[1047]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_10	GGGCGGTGCC	GGCAAGGTCA	CAAAGGCCGC	CGCAAGGCC	TCTAAGAAAT	GA					[1329]
Tcruzi_CLBrener_Esmeraldo_like_1047053511369_20	GGGCGGTGCC	GGCAAGGTCA	CAAAGGCCGC	CGCAAGGCC	TCTAAGAAAT	GA					[1350]
Leishmania_major_F17_0080	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Lmajor_F17_0081	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Lmajor_F17_0082	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Lmajor_F17_0083	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Lmajor_F17_0084	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Lmajor_F17_0085	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Lmajor_F17_0086	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Linfantum_J17_V3_0090	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Linfantum_J17_V3_0100	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Linfantum_J17_V3_0110	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1349]
Linfantum_J17_V3_0170	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Linfantum_J17_V3_0180	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Linfantum_J17_V3_0190	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]
Linfantum_J17_V3_0200	GGGCAGCGGC	GGTAAGGTGA	CAAAGGCCGC	CGCAAGGCT	TCGAAGAAAT	AA					[1350]

Figure 4.3. RAxML ML tree with bootstrap support to study the evolution of EF1 α in the euglenids. The left or top value is the RAxML (ML) bootstrap score (bs) and the MrBayes (MLE) posterior probability (pp) is below or to the right. Only bs $\geq 50\%$ and pp ≥ 0.80 are shown here and the dark bars represent a node with $\geq 95\%$ bs and 1.00 pp recovered from both the ML and MLE analyses, respectively. A single dash represents that that relationship was not recovered in that particular analysis type and a double dash represents that the relationship was recovered, but it was not supported at the thresholds stated. Genus and clade names are to the right even if they were not recovered as monophyletic, with letters next to them to distinguish them for discussion purposes. Heteronematina is labeled at the base of the tree because the members of this group are recovered as basal to the Euglenea in prior studies, even though the group is not monophyletic. The Heteronematina OTU which is recovered within the Euglenea rather than with the within the Heteronematina group is labeled with an asterisk. When a clade consisting of more than one OTU is recovered together as they were in previous literature findings, then the taxa are shaded gray. Hash marks on a branch denote it is represented at one half the actual length. A yellow outline surrounds the OTUs with divergent sequences that comprise the “crown group”.



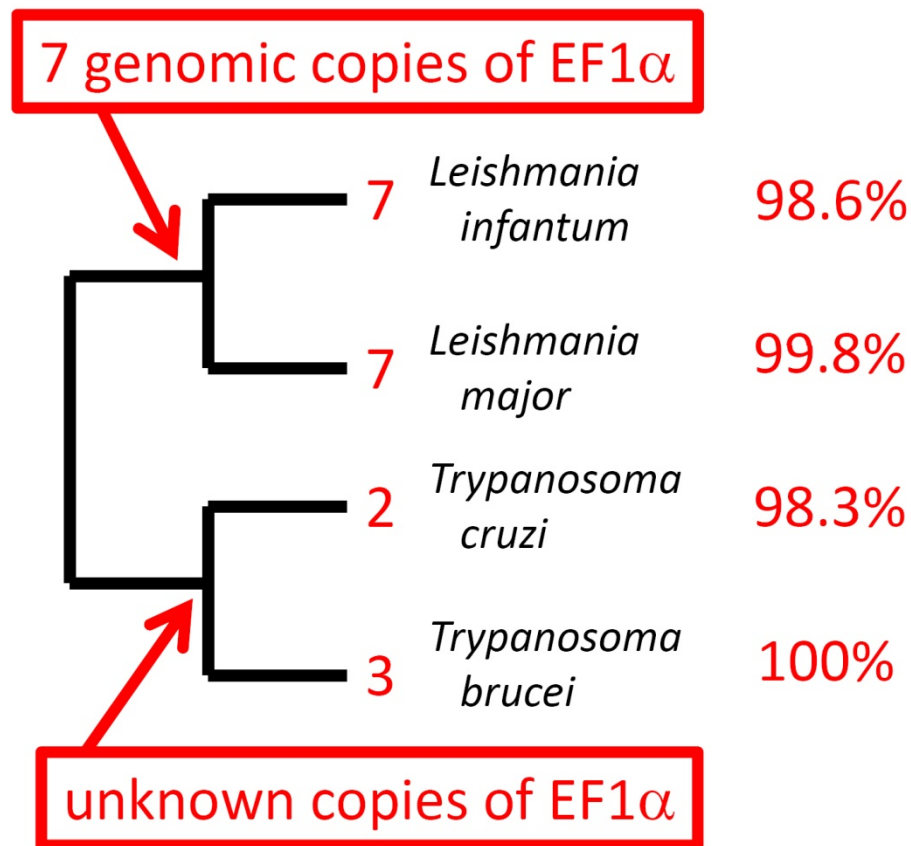


Figure 4.4. Cartoon of the RAxML ML tree used to evaluate the evolution of EF1 α in the kinetoplastids. Each major branch shown in this cartoon was supported by 100% bootstrap support in the original tree. OrthoMCL recovered the number of paralogs (red) in each species that is given after each tip, which is followed by the species name. To the right of the species name is the total percent (red) of identical sites shared between the paralogs in that species. The ancestral genome copy number of EF1 α is commented upon at each node in red.