GENOMIC ANALYSIS OF HISTOMONAS MELEAGRIDIS

by

MICHAEL EDWARD KLODNICKI

(Under the Direction of Robert B. Beckstead)

ABSTRACT

Histomonas meleagridis, a flagellated protozoan of the Order Trichomonadida, is the causative agent of blackhead disease in gallinaceous birds. Few genes have been identified in this organism; thus, little is known regarding the molecular basis for metabolism, virulence, and antigenicity of this parasite. To identify new genes, a cDNA library derived from a lab strain of *H. meleagridis* was sequenced and annotated. Data obtained from these experiments identified 3,425 H. meleagridis genes. Analysis of the data allowed the identification of 81 genes coding for putative hydrogenosomal proteins and was used to determine the codon usage frequency. Sequence information also identified bacteria that are cultured with *H. meleagridis*. Future analysis of these data should provide valuable molecular insights into *H. meleagridis* and provide the platform for molecular studies aimed at understanding the pathogenesis of blackhead disease.

INDEX WORDS: *Histomonas meleagridis*, blackhead, hydrogenosome

GENOMIC ANALYSIS

OF HISTOMONAS MELEAGRIDIS

by

MICHAEL EDWARD KLODNICKI

B.Sc., University of Georgia 2012

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

MASTER OF SCIENCE

ATHENS, GEORGIA

2013

GENOMIC ANALYSIS

OF HISTOMONAS MELEAGRIDIS

by

MICHAEL EDWARD KLODNICKI

Major Professor: Robert B. Beckstead

Committee: Larry R. McDougald

Gene Pesti

Electronic Version Approved:

Maureen Grasso

Dean of the Graduate School

The University of Georgia

December 2013

DEDICATION

I would like to dedicate this thesis to the domesticated turkey, in the hopes that they will suffer from *Histomonas* no more.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Dr. Robert B. Beckstead, my major professor, for his guidance, support and encouragement throughout my undergraduate and graduate student careers at The University of Georgia. I would like to also thank my committee members, as well as Dr. Lorraine Fuller, Dr. John Shields, Dr. William Kerr, and Dr. Mark Haidekker for technical assistance during my study.

TABLE OF CONTENTS

Page

CHAPTER 1
INTRODUCTION AND LITERATURE REVIEW

OVERVIEW

Histomonas meleagridis, an early-diverging unicellular eukaryote, infects a wide range of gallinaceous birds, causing histomoniasis (blackhead disease) (McDougald 2005). This disease is of commercial importance, sometimes causing high morbidity and mortality in turkeys and peafowl. Little is known of the molecular biology of this protozoan, although much can be inferred from work on closely related members of the Tritrichomonadidae. All species of this order are anaerobic and lack mitochondria, depending instead on the hydrogenosome for glycolytic energy metabolism and formation of molecular hydrogen (Lindmark and Muller 1973).

Genes involved in virulence and metabolism have been studied extensively in the human parasite *Trichomonas vaginalis*, the cattle parasite *Tritrichomonas foetus*, and in a more distantly related human parasite, *Entamoeba histolytica* (Alderete et al., 1991a, b; Addis et al., 1999; Musatovova and Alderete, 1999; Chaudhry and Petri, 2005; Mundodi et al., 2006, 2008). Given the morphologic and physiological similarity and phylogenetic relationship of *H. meleagridis* to these organisms, it is assumed that many genes are held in common. However, the amount of specific sequence data in *H. meleagridis* remains limited.

TRANSMISSION AND PATHOLOGY

H. meleagridis is a protozoan parasite with a complex life cycle, and one that requires an intermediate host for transmission. *Heterakis gallinarum*, a cecal-dwelling nematode, is the key carrier of this parasite – with *H. gallinarum* itself often carried by earthworms. Histomonads are shed by ingested *H. gallinarum* during its first molt, in the gut of the bird; these shed parasites then migrate to the bird's cecal wall, which becomes thickened with inflammation, and characterizes the site of primary infection. Parasites further migrate into the blood, and invade the liver via connecting portal blood vessels, establishing a secondary site of infection (Clarkson 1961; Fine 1975). These infections produce characteristic lesions in the ceca and the liver, with only the former correlated with bacterial infections (Franker et al. 1964; Bradley and Reid 1966). *Escherichia coli, Clostridium perfringens or Bacillus subtilis* typically cause these secondary bacterial infections associated with mortality.

Transmission from bird to bird, or lateral transmission, has been observed and explained by cloacal drinking, or the bird's reflexive, reverse-peristaltic cloacal movement during contact with infected fecal matter. Thus, for *in vivo* studies of histomoniasis, birds are often experimentally infected via the cloacal drop method (Hu et al 2004). Furthermore, chickens and other birds which may not show clinical signs of *Histomonas* infection can still act as carriers of infected *H. gallinarum* eggs, which can remain viable and infected with virulent strains of *H. meleagridis,* with periods of up to three years reported.

PREVENTION AND TREATMENTS

Various nitroheterocyclics, particularly nitroimidazoles like dimetridazole, were used for many years and greatly reduced the incidence of blackhead disease. The FDA and the EU banned use of these compounds in the 90s and in 2003, respectively, as the imidazoles were suspected carcinogens. Nitarsone is currently the only drug approved for use in animals in the prevention of blackhead disease (Hu and McDougald 2004). Recommended for use as a feed additive at 187.5 ppm, it reduces flock mortality but allows relapse upon withdrawal of treatment (McGuire and Morehouse 1952). Banned in the EU, nitarsone is the last arsenic-containing drug on the US market, with the FDA recently facing pressure to withdraw its approval of nitarsone use in animals bound for human consumption in light of concerns raised by recent studies (Nachman et al 2013).

Alternative to preventative administration of arsenic-containing compounds is vector control, or use of antihelminthics, such as benzimidazoles, which must be administered prior to vector exposure, and do not provide protection from lateral transmission or the protozoa itself (Hegngi 1999, Hu et. al 2004). As of late 2013, there is no treatment available for histomoniasis.

IN VITRO CULTURE

H. meleagridis can be isolated *in vitro* from the infected ceca and liver using a Dwyer medium (McDougald and Galloway 1973). Formulation of Dwyer's media consisted of 85-95% of Medium 199, 5-10% horse serum, 5% chick embryo extract, and 1% rice powder (Dwyer 1970), but was recently modified by removing the chick embryo extract and reducing the rice flour concentration to 0.8% (van der Heijden and Landman 2007). This modified version of Dwyer's medium facilitates rapid expansive growth of both *H. meleagridis* and bacteria and is commonly used by researchers today.

Histomonads must be cultured in slightly acidic media under strict anaerobic conditions. The typical passage time of cultures varies from 48 to 72 hours, and peak growth is obtained in 1-5 days. Virulence of monad strains appears to be roughly inversely proportional to number of passages (Dwyer 1970, Lund et. al 1967)

Histomonads must be co-cultured with bacteria to maintain cultures over time. However, Lund et al. (1967) reported several experiments in which antibiotic-treated bacteria were used to support culture growth. Lesser (1960a, 1960b, 1963, 1964) attempted to remove serum from the growth medium by substitution with sterilized bovine cream and various fatty acids supplemented with various mineral salts; he used uncharacterized antibiotic-treated bacteria extracted from turkey cecal samples to maintain cultures. Additionally, he successfully used antibiotic-treated tissues from fresh hamster liver in lieu of bacteria for prolonged (more than 190) passages of histomonad cultures.

Other studies in closely related cattle parasites (trichomonads) demonstrated inhibitory effect of zinc treatment on *Tritrichomonas foetus* growth *in vitro* (Benchimol et al 1993). Similarly, nitarsone was demonstrated effective in controlling the growth of *H. meleagridis* in a dose dependent manner (Van der Heijden et. al, 2008).

STUDY OBJECTIVES

1. Sequence and annotate a cDNA library, to provide data for future molecular and virulence studies.
2. Analyze transcriptomic data, and present potential unique molecular targets for therapeutic treatment of histomoniasis.

LITERATURE CITED

Addis, M. F., P. Rappelli, A. M. Pinto De Andrade, F. M. Rita, M. M. Colombo, P. Cappuccinelli, and P. L. Fiori. 1999. Identification of Trichomonas vaginalis alpha-actinin as the most common immunogen recognized by sera of women exposed to the parasite. The Journal of Infectious Diseases 180: 1727-1730.

Alderete, J. F., E. Newton, C. Dennis, and K. A. Neale. 1991a. Antibody in sera of patients infected with Trichomonas vaginalis is to trichomonad proteinases. Genitourinary Medicine 67: 331-334.

Alderete, J. F., E. Newton, C. Dennis, and K. A. Neale. 1991b. The vagina of women infected with Trichomonas vaginalis has numerous proteinases and antibody to trichomonad proteinases. Genitourinary Medicine 67: 469-474.

Aoyagi, S., and D.H. Baker. 1995. Effect of high copper dosing on hemicellulose digestibility in cecectomized cockerels. Poultry Sci. 74:208-211.

Benchimol, M., Almeida, J.C.A., Lins, U., Goncalves, N.R., Souza, W. 1993. Electron microscopic study of the effect of zinc on *Tritrichomonas foetus*. Microb. Agents Chem. 37:2722-2726.

Bradley, R. E., and W. M. Reid. 1966. *Histomonas meleagridis* and several bacteria as agents of infectious enterohepatitis in gnotobiotic turkeys. Exp.Parasitol. 19:91-101.

Chaudhry, O. A., and W. A. Petri. 2005. Vaccine prospects for amebiasis. Expert Review of Vaccines 4: 657-668.

Clarkson, M. J. 1961.The blood supply of the liver of the turkey and the anatomy of the biliary tract with reference to infection with *Histomonas meleagridis*. Res. Vet. Sci. 2:259-264.

Duke, G.E., and Bedbury, H.P. 1983. Cecal microflora of turkeys fed low or high fiber diets: enumeration, identification, and determination of cellulolytic activity. Poultry Sci. 62(4):675-82.

Duke G.E., Eccleston, E., Kirkwood, S., Louis, C.F., and Bedbury, H.P. 1984. Cellulose digestion by domestic turkeys fed low or high fiber diets. J Nutr. 114(1):95-102.

Dwyer, D. M. 1970. An improved method for cultivating *Histomonas meleagridis*. J Parasitol, *56*(1), 191-192.

Fine, P. E. M. 1975.Quantitative studies on transmission of *Parahistomonas wenrichi* by ova of *Heterakis gallinarum*. Parasitology 70:407-417.

Franker, C.K. and Doll, J.P. 1964. Experimental histomoniasis in gnotobiotic turkeys: effects of some cecal bacteria on pathogenesis. J Parasitol. 50:636-40.

Harms, R.H., and Buresh, R.E. 1987. Influence of three levels of copper on the performance of turkey poults with diets containing two sources of methionine. Poult Sci. 1987 Apr;66(4):721-4.

Hegngi, F.N., Doerr, J., Cummings, T.S., Schwartz, R.D., Saunders, G., Zajac, A., Larsen, C.T., and F.W. Pierson. 1999. The effectiveness of benzimidazole derivatives for the treatment and prevention of histomonosis (blackhead) in turkeys. Vet Parasitol 81: 29-37.

Hu, J. & McDougald, L.R. 2004. The efficacy of some drugs with known antiprotozoal activity against *Histomonas meleagridis* in chickens. Veterinary Parasitology, 121, 233-238.

Hu, J., Fuller, L., and L. R. McDougald. 2004. Infection of turkeys with *Histomonas meleagridis* by the cloacal drop method. Avian Dis 48: 746-750.

King, J.O.L. 1975a. The feeding of copper sulphate to growing rabbits. Br. Vet. J. 131:70–75.

King, J.O.L. 1975b. The feeding of copper sulphate to ducklings. Br. Poult. Sci. 16:409–411.

Leeson S., Zubair, A.K., Squires, E.J., and Forsberg, C. 1997. Influence of dietary levels of fat, fiber, and copper sulfate and fat rancidity on cecal activity in the growing turkey. Poult Sci. 76(1):59-66.

Lesser, E. 1961. *In vitro* cultivation of Histomonas meleagridis free of demonstrable bacteria. Journal of Protozoology 8: 228-230.

Lesser, E. 1963. Effect of amphotericin-B on in vitro growth of Histomonas meleagridis. Journal of Parasitology 49: 329.

Lesser, E. 1964. I. Studies on the in vitro cultivation of Histomonas meleagridis with three different mixed bacterial flora. II. Studies on the in vitro growth of Histomonas meleagridis with single species of bacteria. Proc. Helminth. Soc. Wash. 31: 263-264; 265-266.

Lindmark, D. G., and M. Muller. 1973. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate, *Tritrichomonas foetus,* and its role in pyruvate metabolism. Journal of Biological Chemistry 248: 7724-7728.

Lund, E. E., Augustine, P.C., and Chute, A.M. 1967. *Histomonas meleagridis* after one thousand in vitro passages. Journal of Protozoology 14: 349–351.

Lund, E. E., Wehr, E.E., Ellis, D.J. 1966. Earthworm Transmission of *Heterakis* and *Histomonas* to turkeys and chickens. The Journal of Parasitology 52: 899-902.

McDougald, L. R., and R. B. Galloway. 1973. Blackhead disease: in vitro isolation of *Histomonas meleagridis* as a potentially useful diagnostic aid. Avian Dis. 17:847-50.

Mcdougald, L. R. 2005. Blackhead disease (histomoniasis) in poultry: a critical review. Avian Dis*, 49*(4), 462-476.

McGuire, W. C., and N. F. Morehouse. 1952. Chemotherapy studies of histomoniasis. Poult. Sci. 31:603-609.

Mundodi, V., A. S. Kucknoor, and J. F. Alderete. 2008. Immunogenic and plasminogen-binding surface-associated alpha-enolase of *Trichomonas vaginalis*. Infection and Immunity 76: 523-531.

Mundodi, V., A. S. Kucknoor, T. H. Chang, and J. F. Alderete. 2006. A novel surface protein of *Trichomonas vaginalis* is regulated independently by low iron and contact with vaginal epithelial cells. BMC Microbiology 6: 6.

Musatovova, O., and J. F. Alderete. 1999. The *Trichomonas vaginalis* phenotypically varying P270 immunogen is highly conserved except for numbers of repeated elements. Microbial Pathogenesis 27: 93-104.

Nachman, K.E., Baron, P.A., Raber, G., Francesconi, K.A., Navas-Acien, A., and Love, D.C. 2013. Roxarsone, inorganic arsenic, and other arsenic species in chicken: a U.S.-based market basket sample. Environ Health Perspect. 121(7):818-824.

Persia, M.E., Baker, D.H., Parsons, C.M. 2004. Tolerance for excess basic zinc chloride and basic copper chloride in chicks. Br Poult Sci. 45(5):672-676.

Rolls, B.A. 1977. Inorganic ions in the intestinal and cecal contents of germ-free and conventional chickens. Lab. Anim. 11:99-104.

van der Heijden, H. M. J. F., and W. J. M. Landman. 2007. Improved culture of *Histomonas meleagridis* in a modification of dwyer medium. Avian Dis 51: 986-988.

van der Heijden H.M., and Landman W.J. 2008. *In vivo* effect of herbal products against *Histomonas meleagridis* in turkeys. Avian Pathol. 37(1):45-50.

CHAPTER 2

A GENOMIC ANALYSIS OF *HISTOMONAS MELEAGRIDIS* THROUGH SEQUENCING OF A CDNA LIBRARY

M. Klodnicki, L. R. McDougald and R. B. Beckstead*

Accepted by *Journal of Parasitology*. Reprinted here with permission of Allen Press.

ABSTRACT

Histomonas meleagridis, a flagellated protozoan of the Order Trichomonadida, is the causative agent of blackhead disease in gallinaceous birds. Few genes have been identified in this organism; thus, little is known regarding the molecular basis for metabolism, virulence, and antigenicity of this parasite. To identify new genes, a cDNA library derived from a lab strain of *H. meleagridis* was sequenced and annotated. Data obtained from these experiments identified 3,425 *H. meleagridis* genes. Analysis of the data allowed the identification of 81 genes coding for putative hydrogenosomal proteins and was used to determine the codon usage frequency. Sequence information also identified bacteria that are cultured with *H. meleagridis*. Future analysis of these data should provide valuable molecular insights into *H. meleagridis* and provide the platform for molecular studies aimed at understanding the pathogenesis of blackhead disease.

INTRODUCTION

Histomonas meleagridis, an early-diverging unicellular eukaryote, infects a wide range of gallinaceous birds, causing histomoniasis (blackhead disease) (McDougald, 2005). This disease is of commercial importance, sometimes causing high morbidity and mortality in turkeys and other birds. Little is known of the molecular biology of this protozoan, although much can be inferred from work on closely related members of the Tritrichomonadidae. All species of this order are anaerobic, lacking mitochondria**,** but depending instead on the hydrogenosome for glycolytic energy metabolism and formation of molecular hydrogen (Lindmark and Muller, 1973).

Genes involved in virulence and metabolism have been studied extensively in the human parasite *Trichomonas vaginalis*, the cattle parasite *Tritrichomonas foetus*, and

in a more distantly related human parasite, *Entamoeba histolytica* (Alderete et al., 1991a, b; Addis et al., 1999; Musatovova and Alderete, 1999; Chaudhry and Petri, 2005; Mundodi et al., 2006, 2008). Given the morphologic and physiological similarity and phylogenetic relationship of *H. meleagridis* to these organisms, it is assumed that many genes are held in common. However, the amount of specific sequence data in *H. meleagridis* remains limited.

In the present study, we generated, sequenced, and annotated a cDNA library from a lab culture of *H. meleagridis*. Using this information, 3,425 putative *H. meleagridis* genes were identified, a detailed characterization of hydrogenosomal proteins was carried out, and the codon preference was determined.

MATERIALS AND METHODS

An isolate of *H. meleagridis* obtained from a backyard poultry flock located in Buford, Georgia was passed in vitro in modified Dwyer's medium containing 89% M199, 10% non-sterile deactivated horse serum, 0.1% sodium bicarbonate, and 1% rice powder (Dwyer, 1970; Hauck et al., 2010). Cultures were grown at 40° C. Total RNA was isolated from *H. meleagridis* using TRIzol (Life Technologies, Carlsbad, California) according to manufacture protocol. cDNA was produced using the MINT-Universal cDNA synthesis kit using an oligo-dt primer (Evrogen Joint Stock Company, Moscow, Russia) to specifically detect and bind poly-A tails of transcripts in order to reduce bacterial transcript contamination. cDNA output was normalized using the TRIMMER cDNA normalization kit (Evrogen Joint Stock Company). The cDNA library was sequenced using the Roche 454 platform at the Georgia Genomics Facility, Athens, Georgia. Sequence reads were aligned to generate continuous sequence reads (contigs)

using the Roche GS de novo Assembler under standard settings (Roche, Basle, Switzerland).

Analyses of contig sequences were made using the BLASTx and BLASTn algorithms (Altschul et al., 1990) and the BLAST2GO application (Conesa et al., 2005) for local alignment comparisons to the NCBI database of genomic information, GenBank (Benson et al., 2011) using BLAST2GO default parameters (maximum e-value of 1.0e-3). Database entries were parsed for redundancy (entries with greater than 95% invariance were determined to be variants of the same gene) and compared to cDNA contig entries using basic local pairwise alignment analyses. Intensive local alignments and gene constructs were generated and evaluated using the BLOSUM62 algorithm native to the application Geneious (Drummond et al, 2010). Positively identified sequences were submitted online to Genbank via BankIt.

BLAST sequence alignments were filtered by organism**;** protozoan and bacterial hits were assembled into sets and examined separately. Bacterial hits were filtered by phylogenetic family and the number of hits belonging to the 10 most frequent-occurring families reported. For analyses of the protozoan hits, sequence Manipulation Suite 2, or SMS2 (Stothard, 2000) was used to process and identify open reading frames (ORFs) and for further sequence manipulation and examination, including the generation of predicted codon usage data. Using SMS2, a sample of 100 of the longest ORFs in the *H. meleagridis* data set was generated for codon usage data mining. *Histomonas meleagridis* putative hydrogenosomal proteins were identified by homology to known proteins found in *T. vaginalis* (Henze, 2007). Contigs were evaluated for redundancy by local nucleotide alignments.

RESULTS

Sequencing of the *H. meleagridis* cDNA library resulted in approximately 24MB of DNA sequences contained in 70,000 unique sequence reads. Assembly of the sequence reads yielded a total of 5,711 contiguous DNA sequences (contigs) of 150 base pairs or greater (Table I), with an average length of 919 base pairs per contig. Blastx analysis of the contigs resulted in 5,430 sequences showing homology to a protein in the NCBI database and 281 contigs without blast hits. The majority of the 281 contigs had 1 or more short open reading frames, suggesting that they may be the result of faulty sequence alignment. Analysis of the sequence database revealed 4,924 discrete protozoan BLASTx hits. Due to the bacteria necessarily present in culture with the monads, analysis also revealed 383 total bacterial hits, which corresponded to 338 unique genes. From the sequence that showed homology to protozoan genes, 74 contigs showed greatest homology to *H. meleagridis* sequences in the NCBI database; these sequences corresponded to 42 unique genes. Parsing the NCBI protein database for a nonredundant list of *H. meleagridis* sequences, it was determined there are 44 unique coding genes corresponding to 85 deposited sequences. Only 2 relatively short genes previously reported from *H. meleagridis* (putative 60S ribosomal protein P1, with a predicted length of 41 amino acid residues; putative histone 2A-IV, 74 predicted residues) were not identified in the cDNA library.

Of the 5,430 total contigs that returned BLAST hits, 4,537 had identified a *T. vaginalis* gene with the greatest respective homology (Fig. 1), suggesting that the *H. meleagridis* transcriptome is more closely related to *T. vaginalis* than other sequenced protozoan genomes. Significant numbers of hits from other related protozoan species, particularly *H. meleagridis* and *E. histolytica,* are also present among the top-hit

distribution. Insignificant numbers of random hits from various, less closely-related species are grouped into the 'Other' category, composing the remainder of top-hits within the protozoan category.

As *Histomonas meleagridis* growth in culture requires bacteria that are co-cultured from the caecal content at the time of isolation, total RNA isolation from *H. meleagridis* cultures results in detection of both bacterial and *H. meleagridis* RNA. Based on the number of bacterial contigs identified in the library, we had 7.3% contamination. Analysis of the bacterial contigs filtered by phylogenetic family allowed the identification of bacteria found cultured with *H. meleagridis*. The majority of bacterial sequences were classified as part of the Clostridiaceae family (obligate anaerobes), the Bacteriodaceae family (aerotolerant anaerobes) or the *Bacillus* genus of the Bacillaceae family (facultative or obligate anaerobes) (Fig. 2).

mRNAs coding for putative hydrogenosomal proteins (Table II) were assembled based on positive experimental evidence for hydrogenosomal localization in trichomonads and data directly resulting from the first proteomics study of the hydrogenosome of *T. vaginalis* (Henze et al., 2007). Local nucleotide alignment analysis of contigs with similar or identical percent identity resulted in 81 unique contigs with significant homology to known hydrogenosomal proteins, composing 40 different enzyme functions. The percent identity of these proteins, as defined by the ratio of positive amino acid matches to the total length of the alignment, ranged from 43% up to 100%. BLAST hits returned e-values ranging from 0 to 1.94e-02. Notably, all enzymes required for ATP synthesis were identified, as were orthologues of 3 confirmed hydrogenosomal membrane proteins (Hmp31 MCF, Hmp35, and TIM23). A number of the proteins identified in the initial *T. vaginalis* proteomics study, however, were

missing, including Hsp10 and Hsp20 (both localized to the hydrogenosome in *T. vaginalis*), along with arginine deiminase, glyoxylase, several chaperones, and several predicted proteins that have not been confirmed to be localized or active within the hydrogenosome. In total, 51 putative hydrogenosomal proteins out of 85 identified were missing from the dataset.

We analyzed the nucleotide and codon usage frequencies of *H. meleagridis* using the longest 100 open reading frames consisting of 57,059 codons (Fig. 3). A strong A/T bias (68%) was observed in the coding sequences. Codon usage in the first position was biased towards purines (62%), while the second and third positions had a higher A/T frequency (65.8% and 66.8%, respectively). Several amino acids exhibited a strong preference for a specific codon, i.e., serine residues existed primarily as TCA codons (38%), proline as CAA codons (63%), threonine as AAA codons (53%), glutamic acid as GGA codons (79%), and aspartate was coded for by the GGT codon (70%). All 3 stop codons were used, with TAA having the highest frequency (64%).

DISCUSSION

The present work has greatly expanded the list of genes known from *H. meleagridis* by comparison of a sequenced cDNA library with genes reported from other organisms. However, it is important to emphasize that many other genes could be discovered through similar studies of RNA isolated from other stages of growth. For instance, as the parasite moves from the anaerobic cecal environment to the liver and other oxygen-rich tissues, it is likely that other genes will be expressed. Ultimately, the sequencing of the entire genome will be needed to identify the complete genetic complement of this organism.

While Hauck and Hafez (2010) found *H. meleagridis* to be more closely related to *Tritrichomonas foetus* based on phylogenetic rRNA studies, limited *T. foetus* sequencing data, relative to the completely sequenced *T. vaginalis* genome, likely prevented the majority of hits from identifying as *T. foetus* homologues. Four genes, most notably superoxide dismutase (Table II) did indeed identify with *T. foetus* (Fig. 1), yet the overwhelming majority of genes identified clearly shared primary homology with *T. vaginalis*. The high level of homology of *H. meleagridis* with *T. vaginalis* emphasizes the close physiological, metabolic, and phylogenetic relationship of the 2 organisms. Thus, it is reasonable to believe that similar mechanisms of virulence and antigenicity can also be found. Identification of all the key ATP-producing enzymes identified and many more putative hydrogenosomal enzymes in the present work provides a general outline for future proteomic studies of the hydrogenosome of this species. In the first proteomics study of *T. vaginalis* (Henze et al., 2007), the predictive target signal peptide data were used extensively alongside proteome analysis and experimental evidence, yielding a large number of putative hydrogenosomal proteins that remained unconfirmed to be a part of the hydrogenosomal proteome. Furthermore, only 30 proteins currently have been demonstrated to localize to the organelle (Burstein et al., 2012). Thus, it is expected that some of the hydrogenosomal genes in *H. meleagridis* were not completely elucidated, such as the signal peptidase that shared homology with *E. histolytica* rather than *T. vaginalis*. In contrast, two different classes of heat shock proteins, Hsp10 and Hsp20, were notably missing from the dataset, yet 8 unique Hsp70 homologs were present (data not shown), only 1 of which was annotated as localized in the hydrogenosome. Fourteen unique thioredoxin family proteins and 3 thioredoxin reductase homologs were also identified. Current and future hydrogenosomal

proteomics studies of trichomonads (Schneider et al., 2011; Burstein et al., 2012) will facilitate further comparative studies of the hydrogenosome of *H. meleagridis*.

As the histomonads for this study were grown in cultures heavily populated with bacteria, and such bacteria are considered the primary food source for culture-grown histomonads, the identification of bacterial genes was expected. Numerous bacteria can be seen in food vacuoles by microscopic examination of histomonads (Mazet et al., 2008), and it is well known that the cultured cecal forms of *H. meleagridis* are intolerant to oxygen (Hauck et al., 2010). As shown in Figure 2, the majority of bacterial sequences were classified in the Clostridiaceae family (obligate anaerobes), the Bacteriodaceae family (aerotolerant anaerobes), or the *Bacillus* genus of the Bacillaceae family (facultative or obligate anaerobes). While most of the free bacteria were removed by washing during sample preparation, it was possible that some remained. For this reason, oligo-dt primers that specifically detect the poly (A) tail on the mRNA were used for amplification. Since bacterial mRNA does not have a poly (A) tail that will react with oligo-dt primers, this reduced the bacterial contamination of the cDNA library.

The codon usage for *H. meleagridis* and nucleotide preference differs from other parasitic protozoans such as *Leishmania* spp. and *Trypanosoma cruzi*, which have G/C rich genomes, and *T. vaginalis*, which was previously reported to have no nucleotide bias in its coding sequence (Alonso et al., 1992; Langford et al., 1992; and Meade et al., 1997). More recently, however, genome sequencing of *T. vaginalis* (Carlton et al., 2007) revealed a genome-wide nucleotide composition of 67.3% A/T, similar to the 68% A/T composition described in the present study. A/T nucleotide bias of *H. meleagridis* also bears similarity to *Plasmodium* and *Entamoeba* species that have a high percentage of A/T nucleotides in their genomes and in codon position 3 (Char and Farthing, 1992;

Musto et al., 1995). Like *T. vaginalis*, *H. meleagridis* has a bias towards purines in codon position 1 and A/T in position 2. *Histomonas meleagridis* preference for A/T in codon position 3 differs from the preference of *T. vaginalis* for pyrimidine in that position. *Histomonas meleagridis* also differs from *T. vaginalis* in its frequent use of the TGA and TAG as stop codons (Meade et al., 1997). *Trichomonas vaginalis* uses the TAA stop codon 93% of the time, which is consistent with its use of TAAA as the poly (A) signal (Espinosa et al., 2002). The usage of all stop codons suggests that *H. meleagridis* potentially has different poly (A) signals than *T. vaginalis*. Codon usage statistics provided for *H. meleagridis* will impart useful annotations for future molecular, biological, and biochemical studies of the organism.

ACKNOWLEDGMENT

We would like to thank R. Thomas for logistical support on the analysis of the sequence data.

LITERATURE CITED

Addis, M. F., P. Rappelli, A. M. Pinto De Andrade, F. M. Rita, M. M. Colombo, P. Cappuccinelli, and P. L. Fiori. 1999. Identification of *Trichomonas vaginalis* alpha-actinin as the most common immunogen recognized by sera of women exposed to the parasite. The Journal of Infectious Diseases 180: 1727–1730.

Alderete, J. F., E. Newton, C. Dennis, and K. A. Neale. 1991a. Antibody in sera of patients infected with *Trichomonas vaginalis* is to trichomonad proteinases. Genitourinary Medicine 67: 331–334.

Alderete, J. F., E. Newton, C. Dennis, and K. A. Neale. 1991b. The vagina of women infected with *Trichomonas vaginalis* has numerous proteinases and antibody to trichomonad proteinases. Genitourinary Medicine 67: 469–474.

Alonso, G., P. Guevara, and J. L. Ramirez. 1992. Trypanosomatidae codon usage and GC distribution. Memoria Instituo do Oswaldo Cruz 87: 517–523.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers. 2011. GenBank. Nucleic Acids Research 39: 32-37.

Burstein, D., S. B. Gould, V. Zimorski, T. Kloesges, F. Kiosse, P. Major, W. F. Martin, T. Pupko, and T. Dagan. 2012. A machine learning approach to identify hydrogenosomal proteins in *Trichomonas vaginalis*. Eukaryotic Cell 11: 217-228.

Carlton, J. M., R. P. Hirt, J. C. Silva, A. L. Delcher, M. Schatz, Q. Zhao, J. R. Wortman, S. L. Bidwell, U. C. M. Alsmark, S. Besteiro, et al. 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science 315: 207–212.

Char, S., and M. J. G. Farthing. 1992. Codon usage in *Entamoeba histolytica*. International Journal for Parasitology 22: 381–383.

Chaudhry, O. A., and W. A. Petri. 2005. Vaccine prospects for amebiasis. Expert Review of Vaccines 4: 657–668.

Conesa, A., S. Gatz, J. M. Garcia-Gomez, J. Terol, M. Talon, and M. Robles. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674-3676.

Drummond, A. J., B. Ashton, S. Buxton, M. Cheung, A. Cooper, J. Heled, M. Kearse, R. Moir, S. Stones-Havas, S. Sturrock, T. Thierer, and A. Wilson. 2010. Geneious v5.3. http://www.geneious.com

Dwyer, D. M. 1970. An improved method for cultivating *Histomonas meleagridis*. Journal of Parasitology 56: 192-193.

Espinosa, N., R. Hernandez, L. Lopez-Griego, and I. Lopez-Villaseror. 2002. Separable putative polyadenylation and cleavage motifs in *Trichomonas vaginalis* mRNAs. Gene 289: 81-86.

Hauck, R., and H. M. Hafez. 2010. Systematic position of *Histomonas meleagridis* based on four protein genes. Journal of Parasitology 96: 396-400.

Hauck, R., P. L. Armstrong, and L. R. McDougald. 2010. *Histomonas meleagridis*: Analysis of growth requirements in vitro. Journal of Parasitology 96: 1-7.

Henze, K. 2007. The proteome of *T. vaginalis* hydrogenosomes. Microbiology Monographs 9: 163-178. *In* Tachezy J (ed), Hydrogenosomes and mitosomes: mitochondria of anaerobic eukaryotes. Springer, Berlin, Germany.

Langford, C. K., B. Ullman, and S. M. Landfear. 1992. *Leishmania:* Codon utilization of nuclear genes. Experimental Parasitology 74: 360–361.

Lindmark, D. G., and M. Muller. 1973. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate, *Tritrichomonas foetus,* and its role in pyruvate metabolism. Journal of Biological Chemistry 248: 7724-7728.

Meade, J. C., P. H. Shah, and W. B. Lushbaugh. 1997. *Trichomonas vaginalis*: Analysis of Codon Usage. Experimental Parasitology 87: 73–74.

McDougald, L. R. 1998. Intestinal protozoa important to poultry. Poultry Science 77: 1156–1158.

McDougald, L.R. 2005. Blackhead disease (histomoniasis) in poultry: A critical review. Avian Diseases 49: 462–476.

Mundodi, V., A. S. Kucknoor, and J. F. Alderete. 2008. Immunogenic and plasminogen-binding surface-associated alpha-enolase of *Trichomonas vaginalis*. Infection and Immunity 76: 523–531.

Mundodi, V., A. S. Kucknoor, T. H. Chang, and J. F. Alderete. 2006. A novel surface protein of *Trichomonas vaginalis* is regulated independently by low iron and contact with vaginal epithelial cells. BMC Microbiology 6: 6.

Musatovova, O., and J. F. Alderete. 1999. The *Trichomonas vaginalis* phenotypically varying P270 immunogen is highly conserved except for numbers of repeated elements. Microbial Pathogenesis 27: 93–104.

Munsch, M., A. Lotfi, H. M. Hafez, S. Al-Quraishy, and H. Mehlhorn. 2009. Light and transmission electron microscopic studies on trophozoites and cyst-like stages of *Histomonas meleagridis* from cultures. Parasitology Research 104: 683-689.

Musto, H., H. Rodriguez-Maseda, and G. Bernardi. 1995. Compositional properties of nuclear genes from *Plasmodium falciparum*. Gene 152: 127–132.

Schneider, R. E., M. T. Brown, A. M. Shiflett, S. D. Dyall, R. D. Hayes, Y. Xie, J. A. Loo, and P. J. Johnson. 2011. The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. International Journal for Parasitology 41: 1421-34.

Stothard, P. 2000. The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28: 1102-1104.

FIGURES

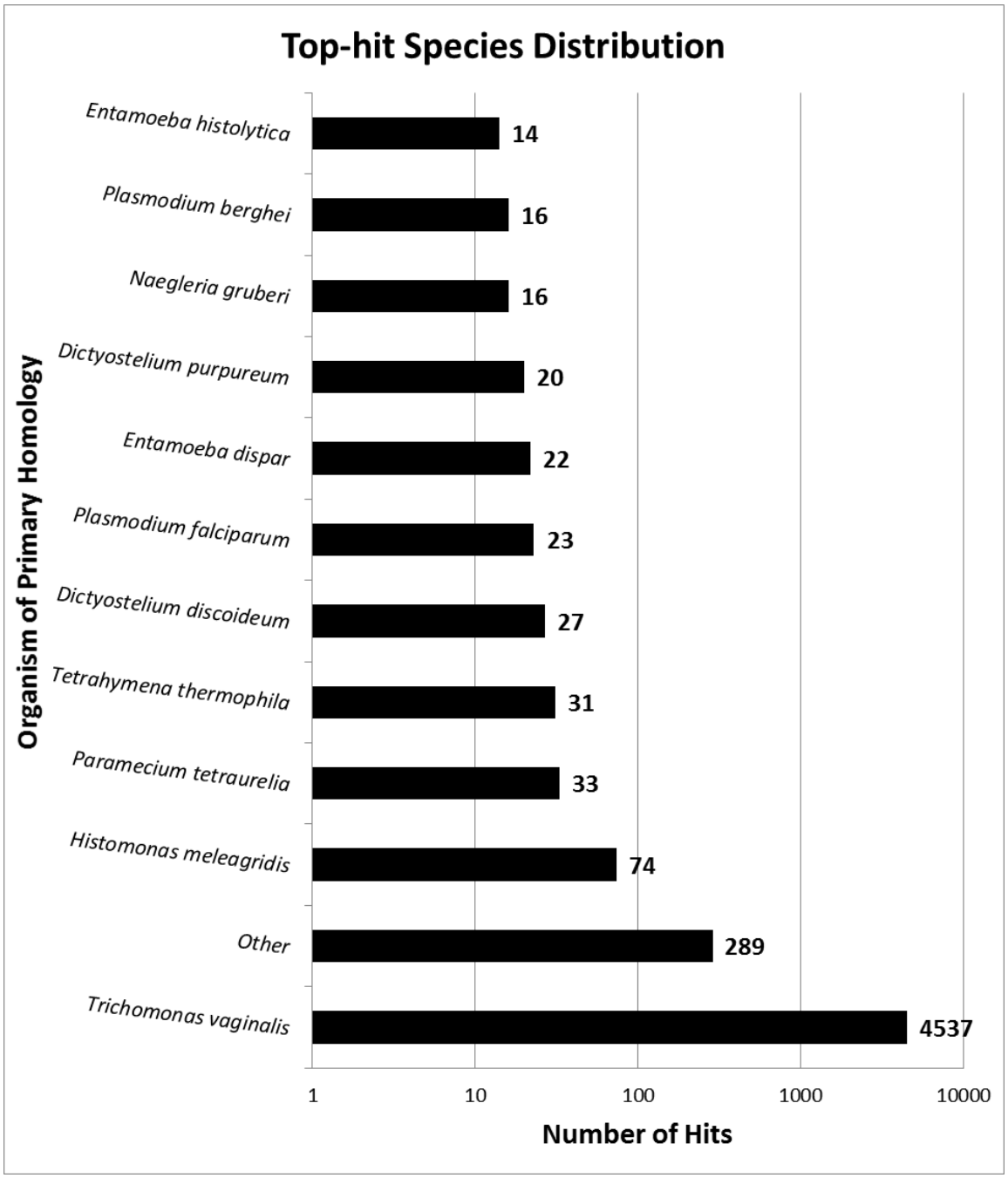

FIGURE 1. Species distribution of total top-homology protozoan hits from BLAST2GO-generated alignments.

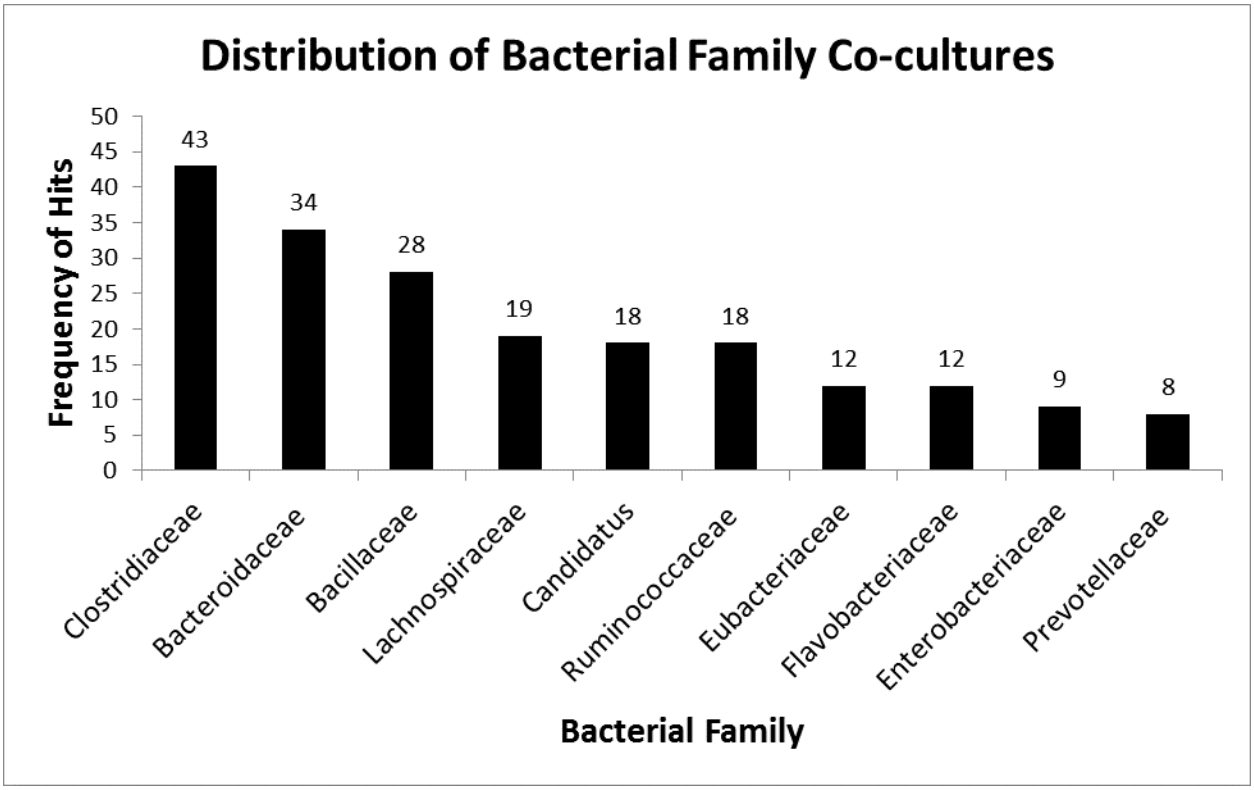

FIGURE 2. Distribution of bacterial species identified in the cDNA contig library. BLAST hits for bacterial species were indexed by phylogenetic family classification and the number of hits belonging to the 10 most frequent-occurring families reported. Frequency indicates the total number of BLAST hits identified within a specific bacterial family. Not shown are 137 bacterial hits belonging to extraneous families.

		Codon Usage Statistics for *H. meleagridis*													
		Second Position													
		T			C			A			G				
			amino			amino			amino			amino			
First Position	T	TTT	Phe	0.48	TCT	Ser	0.18	TAT	Tyr	0.68	TGT	Cys	0.57	T	Third Position
		TTC		0.52	TCC		0.11	TAC		0.32	TGC		0.43	C	
		TTA	LeT	0.36	TCA		0.38	TAA	STOP	0.64	TGA	STOP	0.27	A	
		TTG		0.14	TCG		0.10	TAG		0.09	TGG	Trp	1.00	G	
	C	CTT		0.31	CCT	Pro	0.19	CAT	His	0.71	CGT	Arg	0.35	T	
		CTC		0.12	CCC		0.07	CAC		0.29	CGC		0.11	C	
		CTA		0.04	CCA		0.63	CAA	Gln	0.87	CGA		0.15	A	
		CTG		0.03	CCG		0.10	CAG		0.13	CGG		0.02	G	
	A	ATT	Ile	0.48	ACT	Thr	0.24	AAT	Asn	0.55	AGT	Ser	0.14	T	
		ATC		0.33	ACC		0.12	AAC		0.45	AGC		0.09	C	
		ATA		0.19	ACA		0.53	AAA	Lys	0.53	AGA	Arg	0.27	A	
		ATG	Met	1.00	ACG		0.10	AAG		0.47	AGG		0.09	G	
	G	GTT	Val	0.46	GCT	Ala	0.30	GAT	Asp	0.70	GGT	Gly	0.39	T	
		GTC		0.23	GCC		0.16	GAC		0.30	GGC		0.23	C	
		GTA		0.19	GCA		0.41	GAA	GlT	0.79	GGA		0.33	A	
		GTG		0.11	GCG		0.12	GAG		0.21	GGG		0.06	G	

FIGURE 3. RNA codon usage chart outlining the frequency of codons usage in *Histomonas meleagridis*. Percentages are based on the frequency of the codon in question out of a 100 coding sequences representing 57,059 codons.

Table 2.1. Overview of cDNA library and NCBI statistics.

BLASTx Results	**Total # hits**	**# Unique hits**
5711 contigs (>150bp)	5430	
Bacteria	383	338
Protozoan	4924	3425
T. vaginalis	4537	3152
H. meleagridis	74	42
NCBI Results	**Total # hits**	**# Unique hits**
H. meleagridis	85	44

Table 2.2. Putative hydrogenosomal proteins of *H. meleagridis*.

Homology: function	Accession number	Homology: organism	% Identity to homologous parasites	Homology: accession number
Energy metabolism				
Adenylate kinase	GAAM0100377	*Trichomonas vaginalis*	77%	AAC46483.1
Adenylate kinase family protein	GAAM0101882	*T. vaginalis*	71%	EAY00295.1
	GAAM0103058	*T. vaginalis*	70%	EAY23419.1
Alpha-glucosidase family protein	GAAM0102873	*Bacillus* spp.	77%	CAA54266
	GAAM0103387	*T. thermophila*	43%	EAR96069.1
	GAAM0100154	*T. vaginalis*	79%	EAY00134.1
	GAAM0100172	*T. vaginalis*	79%	EAY00134.1
Alpha-succinyl coenzyme A synthetase	GAAM0100269	*H. meleagridis*	98%	ACI16481
	GAAM0101193	*H. meleagridis*	98%	
	GAAM0101505	*H. meleagridis*	99%	
	GAAM0101957	*H. meleagridis*	98%	
	GAAM0102108	*H. meleagridis*	97%	
Cytosolic malate dehydrogenase 2	GAAM0102023	*H. meleagridis*	93%	CAQ86687
	GAAM0103459	*H. meleagridis*	89%	
Dihydrolipoamide acyltransferase (E2)	GAAM0102527	*T. vaginalis*	69%	EAY01593.1
Glyoxylate reductase	GAAM0101256	*A. thermophila*	66%	BAJ63725.1
Hydrogenosomal malic enzyme subunit A / AP65-2 adhesin	GAAM0103753	*H. meleagridis*	88%	ACI16482
	GAAM0104865		92%	
Hydrogenosomal NADH dehydrogenase 24 kDa subunit	GAAM0104029	*T. vaginalis*	68%	AAV65813.1
Hydrogenosomal NADH dehydrogenase 51-kDa subunit	GAAM0104118	*T. vaginalis*	79%	AAV65812.1
	GAAM0104119	*T. vaginalis*	81%	
	GAAM0104793	*T. vaginalis*	85%	
Hydrogenosomal oxygen reductase	GAAM0103833	*H. meleagridis*	100%	CAQ86684.1
Malate dehydrogenase	GAAM0100034	*T. vaginalis*	77%	EAX95094.1
Oxidoreductase, aldo/keto reductase family protein	GAAM0104093	*T. vaginalis*	65%	EAX97825.1
Oxidoreductase, FAD/FMN-binding family protein	GAAM0101291	*T. vaginalis*	62%	EAY01605.1
	GAAM0103992	*T. vaginalis*	63%	
	GAAM0103997	*T. vaginalis*	63%	
Oxidoreductase, short chain dehydrogenase/reductase family	GAAM0102789	*T. vaginalis*	54%	EAY07632.1
	GAAM0101119	*T. vaginalis*	64%	EAY13790.1
Pyruvate:ferredoxin oxidoreductase D	GAAM0101844	*T. vaginalis*	76%	EAY00742.1
Pyruvate:ferredoxin oxidoreductase E (PFOE)	GAAM0103789	*H. meleagridis*	100%	CAQ86686.1
Succinyl:acetate CoA transferase (ASCT)	GAAM0100115	*G. bemidjiensis*	73%	ACH39847.1
Xanthine dehydrogenase	GAAM0103921	*T. vaginalis*	69%	EAX92968.1
	GAAM0104053	*T. vaginalis*	77%	EAX92968.1
	GAAM0100636	*T. vaginalis*	74%	EAX92968.1
Oxygen scavenging system				
Rubrerythrin	GAAM0102299	*T. vaginalis*	78%	EAY22419.1
	GAAM0102507	*T. vaginalis*	77%	EAY22419.1
Superoxide dismutase 1	GAAM0103633	*Trichomonas foetus*	86%	AAC47734.1
Thiolperoxidase	GAAM0101603	*T. vaginalis*	84%	EAY00427.1
	GAAM0101584	*T. vaginalis*	82%	EAY1083.1
Thioredoxin reductase (txr)	GAAM0100259	*H. meleagridis*	100%	CAQ86691.1
	GAAM0103825	*H. meleagridis*	97%	
	GAAM0100180	*T. vaginalis*	84%	EAY16491.1
Amino acid metabolism				
Aminomethyltransferase	GAAM0105144	*T. vaginalis*	61%	EAY17969.1
Aminotransferase	GAAM0103888	*Clostridium* spp.	84%	EGB93128.1
Fe-S cluster assembly / hydrogenase maturation				
Iron hydrogenase assembly protein	GAAM0103576	*T. vaginalis*	84%	EAY00224.1
Iron hydrogenase	GAAM0100591	*H. meleagridis*	94%	ACI16483
	GAAM0101727	*H. meleagridis*	80%	
Iron-only hydrogenase large subunit	GAAM0101363	*T. vaginalis*	51%	EAY10459.1
NifU-like protein	GAAM0104756	*T. vaginalis*	83%	EAY19669.1

Table 2.2 (continued)

Homology: function	Accession number	Homology: organism	% Identity to homologous parasites	Homology: accession number
Chaperones				
Cpn60	GAAM0100145	*T. vaginalis*	87%	EAX94308.1
	GAAM0100330	*T. vaginalis*	85%	EAX99823.1
	GAAM0105423	*T. vaginalis*	87%	EAX94308.1
	GAAM0105650	*T. vaginalis*	89%	EAX94308.1
DnaJ homolog	GAAM0101518	*T. vaginalis*	84%	EAY09068.1
	GAAM0101492	*T. vaginalis*	65%	EAY16690.1
	GAAM0103557	*T. vaginalis*	53%	EAY17083.1
Hsp70 (hydrogenosomal)	GAAM0102980	*T. vaginalis*	77%	EAY06861.1
SAM50	GAAM0105606	*T. vaginalis*	80%	EAX95236.1
Ubiquitin family protein (UBA/TS-N domain containing protein)	GAAM0101835	*H. meleagridis*	72%	CAQ86697.1
Peptidases				
Cathepsin L-like cysteine protease	GAAM0103559	*T. vaginalis*	75%	EAY00763.1
	GAAM0105441	*T. vaginalis*	75%	EAY01875.1
	GAAM0101540	*T. vaginalis*	65%	EAY08170.1
	GAAM0105301	*T. vaginalis*	65%	EAY08170.1
	GAAM0105564	*T. vaginalis*	57%	EAY08941.1
	GAAM0103911	*T. vaginalis*	56%	EAY09448.1
	GAAM0102472	*T. vaginalis*	60%	EAY11026.1
	GAAM0104634	*T. vaginalis*	76%	EAY13782.1
	GAAM0105690	*T. vaginalis*	78%	EAY13782.1
	GAAM0101698	*T. vaginalis*	71%	EAY16159.1
	GAAM0102958	*T. vaginalis*	56%	EAY15215.1
	GAAM0103856	*T. vaginalis*	62%	EAY06906.1
Signal peptidase	GAAM0105233	*Entamoeba histolytica*	52%	EAL48693.1
Membrane proteins				
Calcium motive P-type ATPase	GAAM0103500	*T. vaginalis*	73%	AAD37691
	GAAM0102036	*T. vaginalis*	48%	EAY16498.1
Hydrogenosomal membrane protein (Hmp31 MCF)	GAAM0100187	*T. vaginalis*	66%	EAY21742.1
	GAAM0100193	*T. vaginalis*	86%	EAY21742.1
	GAAM0100505	*T. vaginalis*	85%	EAY21742.1
Hydrogenosomal membrane protein (Hmp35)	GAAM0101630	*T. vaginalis*	49%	EAX68974.1
TIM23	GAAM0105229	*T. vaginalis*	69%	EAY17723.1

CHAPTER 3
CONCLUSIONS

The present work has greatly expanded the list of genes known from *H. meleagridis* by comparison of a sequenced cDNA library with genes reported from other organisms. However, it is important to emphasize that many other genes could be discovered through similar studies of RNA isolated from other stages of growth. For instance, as the parasite moves from the anaerobic cecal environment to the liver and other oxygen-rich tissues, it is likely that other genes will be expressed. Ultimately, the sequencing of the entire genome will be needed to identify the complete genetic complement of this organism.

While Hauck and Hafez (2010) found *H. meleagridis* to be more closely related to *Tritrichomonas foetus* based on phylogenetic rRNA studies, limited *T. foetus* sequencing data, relative to the completely sequenced *T. vaginalis* genome, likely prevented the majority of hits from identifying as *T. foetus* homologues. Four genes, most notably superoxide dismutase (Table 2.2) did indeed identify with *T. foetus* (Fig. 2.1), yet the overwhelming majority of genes identified clearly shared primary homology with *T. vaginalis*, and not simply because *T. vaginalis* has an assembled genome in reference databases – as evidenced by looking at not just the frequency of top-scoring BLAST hits, but the high percent agreement between alignments. It is illustrative to compare example entries: an adhesin gene with *T. vaginalis* as top-scoring hit may also have *T. foetus* as a secondary hit, but with not as strong a match (lower percent

agreement of alignments). All of this data is now publicly available via NCBI, so the exact phylogenetic relationships of *H. meleagridis* and its protozoan cousins can be more precisely defined with additional sequencing data from these other organisms. Yet, the high level of homology of *H. meleagridis* with *T. vaginalis* emphasizes the close physiological, metabolic, and phylogenetic relationship of these organisms. Thus, it is reasonable to believe that similar mechanisms of virulence and antigenicity can also be found.

Identification of all the key ATP-producing enzymes identified and many more putative hydrogenosomal enzymes in the present work provides a general outline for future proteomic studies of the hydrogenosome of this species. In the first proteomics study of *T. vaginalis* (Henze et al., 2007), the predictive target signal peptide data were used extensively alongside proteome analysis and experimental evidence, yielding a large number of putative hydrogenosomal proteins that remained unconfirmed to be a part of the hydrogenosomal proteome. Furthermore, only 30 proteins currently have been demonstrated to localize to the organelle (Burstein et al., 2012). Thus, it is expected that some of the hydrogenosomal genes in *H. meleagridis* were not completely elucidated, such as the signal peptidase that shared homology with *E. histolytica* rather than *T. vaginalis*. In contrast, two different classes of heat shock proteins, Hsp10 and Hsp20, were notably missing from the dataset, yet 8 unique Hsp70 homologs were present (data not shown), only 1 of which was annotated as localized in the hydrogenosome. Fourteen unique thioredoxin family proteins and 3 thioredoxin reductase homologs were also identified. Current and future hydrogenosomal

proteomics studies of trichomonads (Schneider et al., 2011; Burstein et al., 2012) will facilitate further comparative studies of the hydrogenosome of *H. meleagridis*.

As the histomonads for this study were grown in cultures heavily populated with bacteria, and such bacteria are considered the primary food source for culture-grown histomonads, the identification of bacterial genes was expected. Numerous bacteria can be seen in food vacuoles by microscopic examination of histomonads (Mazet et al., 2008), and it is well known that the cultured cecal forms of *H. meleagridis* are intolerant to oxygen (Hauck et al., 2010). As shown in Figure 2, the majority of bacterial sequences were classified in the Clostridiaceae family (obligate anaerobes), the Bacteriodaceae family (aerotolerant anaerobes), or the *Bacillus* genus of the Bacillaceae family (facultative or obligate anaerobes). While most of the free bacteria were removed by washing during sample preparation, it was possible that some remained. For this reason, oligo-dt primers that specifically detect the poly (A) tail on the mRNA were used for amplification. Since bacterial mRNA does not have a poly (A) tail that will react with oligo-dt primers, this reduced the bacterial contamination of the cDNA library.

The codon usage for *H. meleagridis* and nucleotide preference differs from other parasitic protozoans such as *Leishmania* spp. and *Trypanosoma cruzi*, which have G/C rich genomes, and *T. vaginalis*, which was previously reported to have no nucleotide bias in its coding sequence (Alonso et al., 1992; Langford et al., 1992; and Meade et al., 1997). More recently, however, genome sequencing of *T. vaginalis* (Carlton et al., 2007) revealed a genome-wide nucleotide composition of 67.3% A/T, similar to the 68% A/T composition described in the present study. A/T nucleotide bias of *H. meleagridis* also bears similarity to *Plasmodium* and *Entamoeba* species that have a high percentage of

A/T nucleotides in their genomes and in codon position 3 (Char and Farthing, 1992; Musto et al., 1995). Like *T. vaginalis*, *H. meleagridis* has a bias towards purines in codon position 1 and A/T in position 2. *Histomonas meleagridis* preference for A/T in codon position 3 differs from the preference of *T. vaginalis* for pyrimidine in that position. *Histomonas meleagridis* also differs from *T. vaginalis* in its frequent use of the TGA and TAG as stop codons (Meade et al., 1997). *Trichomonas vaginalis* uses the TAA stop codon 93% of the time, which is consistent with its use of TAAA as the poly (A) signal (Espinosa et al., 2002). The usage of all stop codons suggests that *H. meleagridis* potentially has different poly (A) signals than *T. vaginalis*. Codon usage statistics provided for *H. meleagridis* will impart useful annotations for future molecular, biological, and biochemical studies of the organism.

LITERATURE CITED

Alonso, G., P. Guevara, and J. L. Ramirez. 1992. Trypanosomatidae codon usage and GC distribution. Memoria Instituo do Oswaldo Cruz **87:** 517–523.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers. 2011. GenBank. Nucleic Acids Research **39:** 32-37.

Burstein, D., S. B. Gould, V. Zimorski, T. Kloesges, F. Kiosse, P. Major, W. F. Martin, T. Pupko, and T. Dagan. 2012. A machine learning approach to identify hydrogenosomal proteins in *Trichomonas vaginalis*. Eukaryotic Cell **11**: 217-228.

Carlton, J. M., R. P. Hirt, J. C. Silva, A. L. Delcher, M. Schatz, Q. Zhao, J. R. Wortman, S. L. Bidwell, U. C. M. Alsmark, S. Besteiro, et al. 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science **315:** 207–212.

Char, S., and M. J. G. Farthing. 1992. Codon usage in *Entamoeba histolytica*. International Journal for Parasitology **22:** 381–383.

Espinosa, N., R. Hernandez, L. Lopez-Griego, and I. Lopez-Villaseror. 2002. Separable putative polyadenylation and cleavage motifs in *Trichomonas vaginalis* mRNAs. Gene **289:** 81-86.

Hauck, R., and H. M. Hafez. 2010. Systematic position of *Histomonas meleagridis* based on four protein genes. Journal of Parasitology **96:** 396-400.

Hauck, R., P. L. Armstrong, and L. R. McDougald. 2010. *Histomonas meleagridis*: Analysis of growth requirements in vitro. Journal of Parasitology **96:** 1-7.

Henze, K. 2007. The proteome of *T. vaginalis* hydrogenosomes. Microbiology Monographs **9:** 163-178. *In* Tachezy J (ed), Hydrogenosomes and mitosomes: mitochondria of anaerobic eukaryotes. Springer, Berlin, Germany.

Langford, C. K., B. Ullman, and S. M. Landfear. 1992. *Leishmania:* Codon utilization of nuclear genes. Experimental Parasitology **74:** 360–361.

Mazet, M. M. Diogon, J.F. Alderete, C.P. Vivares, and F. Delbac. 2008. First molecular characterization of hydrogenosomes in the protozoan parasite *Histomonas meleagridis*. International Journal for Parasitology **38:** 177-90.

Meade, J. C., P. H. Shah, and W. B. Lushbaugh. 1997. *Trichomonas vaginalis*: Analysis of Codon Usage. Experimental Parasitology **87:** 73–74.

Musto, H., H. Rodriguez-Maseda, and G. Bernardi. 1995. Compositional properties of nuclear genes from *Plasmodium falciparum*. Gene **152:** 127–132.

Schneider, R. E., M. T. Brown, A. M. Shiflett, S. D. Dyall, R. D. Hayes, Y. Xie, J. A. Loo, and P. J. Johnson. 2011. The *Trichomonas vaginalis* hydrogenosome proteome is

highly reduced relative to mitochondria, yet complex compared with mitosomes. International Journal for Parasitology **41:** 1421-34.