CARBON DIOXIDE CONCENTRATING MECHANISMS IN MARINE DIATOMS: GENETICS, PHYSIOLOGY, AND DIVERSITY

by

CHEN SHEN

(Under the Direction of Brian M Hopkinson)

ABSTRACT

CO_2 concentrating mechanisms (CCMs) have been well studied in model cyanobacteria and in the model green eukaryotic alga, *Chlamydomonas reinhardtii*, while CCMs in marine diatoms are less clearly understood. Given the ecological significance of diatoms, thorough knowledge of diatom CCMs is desired to better understand its role in carbon assimilation and its expected response to future environmental change. In this dissertation, the genetics, physiology and diversity of diatom CCMs are explored using a range of approaches: genetic transformation techniques, Membrane Inlet Mass spectrometry (MIMS), and bioinformatics analyses. First, using genetic transformation techniques, a putative δ-carbonic anhydrase (CA) was localized to the chloroplast periplastidal compartment and three putative HCO_3^- transporters were also successfully localized in the diatom *Thalassiosira pseudonana*, one to the plasma membrane and two to the chloroplast. Physiological assessments were performed to demonstrate the activity and role of an SLC4 (solute carrier 4) HCO_3^- transporter found in the plasma membrane of diatom *Phaeodactylum tricornutum*. The genetic and physiological characterization of CCM components conducted here helps create a more integrated description of CCMs in model diatoms. Secondly, the importance of extracellular carbonic anhydrase (eCA) in supporting CO_2

uptake in centric diatoms was demonstrated by quantitative measurement of eCA activity using MIMS. Assessment of centric diatoms spanning a large size range showed that eCA activity increased with cell radius to support the greater demand of CO_2 for photosynthesis in larger centric diatoms. Photosynthesis was reduced when eCA was inhibited but there was no overall relationship between the effect of eCA-inhibited photosynthesis and cell size. Thirdly, the diversity of marine diatom CCMs was evaluated based on the distribution of CCM components (CA and HCO_3^- transporter) among 34 marine diatom strains whose genome or transcriptome sequences were available. Diatom CCM structure was quite diverse, driven primarily by extensive variation in α-CAs and δ-CAs, indicating radiation of these genes within diatom lineages. Comparing relationships among diatom CCMs (using hierarchical clustering) with diatom species phylogeny (based on an18S rDNA phylogenetic tree) suggests that CCM structure is somewhat congruent with diatom phylogeny.

INDEX WORDS: CO_2 concentrating mechanisms, marine diatoms, carbonic anhydrase, HCO_3^- transporter, sub-cellular localization, Membrane Inlet Mass Spectrometry, bioinformatics

CARBON DIOXIDE CONCENTRATING MECHANISMS IN MARINE DIATOMS: GENETICS, PHYSIOLOGY, AND DIVERSITY

by

CHEN SHEN

BS, Ocean University of China, P. R. China, 2010

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

2016

CARBON DIOXIDE CONCENTRATING MECHANISMS IN MARINE DIATOMS: GENETICS, PHYSIOLOGY, AND DIVERSITY

by

CHEN SHEN

Major Professor:	Brian Hopkinson
Committee:	Brian Binder
	Mary Ann Moran
	Boris Striepen
	Patricia Yager

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2016

DEDICATION

To my supportive parents

ACKNOWLEDGEMENTS

There is no way I would have finished this dissertation or my PhD studies without the help of my advisor, my committee members, other faculty members and staff in Marine Sciences department, and my dear family and friends. My deepest acknowledgements go to my supervisor, Brian Hopkinson. He was willing to answer my trivial questions and always gave positive feedback to my work even when the experiments seemed never work out. His attitude towards sciences teaches me to be persistent and patient and motivates me to stay in sciences. I also want to thank my committee members, Dr. Brian Binder, Dr. Boris Striepen, Dr. Mary Ann Moran and Dr. Patricia Yager for their time and efforts on my work. Their kindness and thoughtful comments made the committee meetings an informative lesson instead of an intimidating exam that I expected them to be. I appreciate that Dr. Mary Ann Moran kindly invited Brian and I to join their lab meeting and allowed me to participate in their paper. I thank other faculty member and staff, your kind help made everything running smoothly during these years. My only labmate Anna, we keep each other accompany and I really enjoyed our field work in Key Largo. I am not lonely in this country because I have my friends, Qian Liu, Yuntao Wang, Linquan Mu, Jing Chen, Yeping Yuan, Bingyu Li, Lu Fan and Valery Flint and my cats. Last but not least, I can’t be more grateful that I have such a supportive and happy family who tell me to work hard and enjoy life.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Motivation

Why study CO_2 Concentrating Mechanisms?

Humans have emitted large amounts of CO_2 into the atmosphere since the beginning of the industrial period as a result of fossil-fuel burning, cement production, agriculture, and deforestation (Sabine et al., 2004). CO_2 is released into the atmosphere where much of it has remained, increasing the partial pressure of CO_2, but both the ocean and the land have taken up a substantial amount of CO_2 released through human activities (Sabine et al., 2004; Riebesell et al., 2007). Covering ~70% of the Earth's surface, the ocean plays an important role in modulating atmospheric carbon dioxide by a variety of physical, chemical, and biological mechanisms (Falkowski et al., 2000; Zeebe and Wolf-Gladrow, 2001; Rost et al., 2008). CO_2 entering the ocean alters the seawater carbonate equilibrium as a result acid-base reactions involving carbon. There is an increase in the total concentration of dissolved inorganic carbon (DIC = CO_2 + HCO_3^- + CO_3^{2-}), carbon dioxide, and bicarbonate ions (HCO_3^-), while there is a decrease in pH and in the concentration of carbonate ions (CO_3^{2-}) (Wolf-Gladrow et al., 1999). This increased acidity of the ocean due to uptake of anthropogenic CO_2 is often referred as 'ocean acidification' (Caldeira and Wickett, 2003). The pH of the world's oceans has already decreased by 0.1 since the industrial revolution began (Caldeira and Wickett, 2003). Assuming a 'business-as-usual emission scenario' (IS92a, Leggett et al., 1992), the atmospheric CO_2 concentration will rise by about a factor of two relative to the present day (~380ppm) by the year 2100, which will result in

seawater pH dropping by a further ~0.3 pH units (Houghton et al., 2001; Riebesell et al., 2007). Ocean acidification alters the inorganic carbon chemistry of seawater and the drop in pH may have numerous chemical and biological consequences, including changes in the availability of nutrients (Millero et al., 2009). In addition, elevated CO_2 and other greenhouse gases will lead to global warming (Broecker, 1975). As the surface ocean warms, ocean mixed layers will on average become shallower and the stratification of the ocean will increase, which will further limit nutrient supply from deep waters (Doney, 2006).

The changes in seawater chemistry and nutrients caused by ocean acidification as well as other environmental problems like global warming are expected to affect phytoplankton at the community and ecosystem level (Wolf-Gladrow et al., 1999; Rost et al., 2008). Importantly, a fraction of the carbon fixed by phytoplankton in the surface ocean subsequently sinks into the ocean interior where it is respired to CO_2 by microbes (Volk and Hoffert, 1985). This process, called the biological pump, is responsible for the sequestration CO_2 in the deep ocean, removing it from the atmosphere. Changes in the strength of the biological pump due to climate change have the potential to alter the ocean's ability to sequester CO_2. For example, changes caused by ocean acidification may differentially affect differently sized phytoplankton, and thereby alter sinking rates and organic carbon export (Flynn et al., 2012; Finkel et al., 2010).

There are multiple mechanisms by which climate change may alter phytoplankton communities and the biogeochemical cycles they catalyze. One of the key variables is thought to be the increase in CO_2 concentrations, which have been found to stimulate growth and photosynthesis in many, though not all, species of phytoplankton (Wolf-Gladrow et al., 1999; Raven et al., 2011). The responses of phytoplankton to increased CO_2 are thought to be mediated by a system know as the CO_2 concentrating mechanism (CCM). Most marine microalgae use a

CCM to overcome the limited availability of CO_2 in seawater and to elevate the concentration of CO_2 around RubisCO, the enzyme involved in the initial fixation of CO_2 in microalgae (Price et al, 2008; Moroney and Ynalvez, 2007; Matsuda et al., 2011). At higher CO_2 concentrations, the CCM is downregulated or even turned off entirely saving energy and materials which can then be put towards increased rates of growth or other metabolism (Kranz et al., 2011). Therefore, it is essential to study CCMs in order to understand the influence of ocean acidification and other climate change variables on phytoplankton and to understand how phytoplankton and the global carbon cycle will respond to these perturbations (Falkowski et al., 2000; Rost et al., 2008; Raven et al., 2011; Reinfelder, 2011).

The Importance of Diatoms

As one of the most ecologically significant primary producers in the ocean, marine planktonic diatoms are responsible for up to 20% of global carbon fixation (Tréguer et al., 1995; Nelson et al., 1995; Field et al., 1998; Mann, 1999). Therefore, the physiological responses of diatoms to elevated CO_2 concentrations are especially important to constrain to better understand changes in ocean ecology and future oceanic carbon sequestration. Diatoms are characterized by their unique and beautiful cell walls made of non-crystalline silica, which display regular geometric structures. The cell wall consists of two large, intricately sculptured parts called valves, with several thinner, linking structures termed girdle elements. The valves and girdle elements fit together tightly so that the flux of material across the wall is limited to pores or slits in the frustule (Round et al., 1990). Diatoms have a rather complicated evolutionary history, including a key secondary endosymbiosis in which an ancestral non-photosynthetic eukaryote engulfed a photosynthetic alga to form the lineage that led to diatoms (Moustafa et al., 2009;

Kroth, 2002). And as a result, diatoms possess a four-layered chloroplastic membrane system (Gibbs, 1981).

1.2 Background

1.2.1 CO_2 concentrating mechanisms

The availability of dissolved CO_2 in the seawater is limited since the diffusion rate of dissolved CO_2 in water is very slow compared to that of gaseous CO_2 in air, and the CO_2-HCO_3^- equilibrium favors HCO_3^- as a result of the high pH and salinity of seawater (Goyet and Poisson, 1989; Raven et al., 1994; Matsuda et al., 2001). Moreover, the low affinity of Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), the principal enzyme that catalyzes carbon fixation in the Calvin-Benson cycle for dissolved CO_2 is not sufficient to sustain high photosynthetic rate under current oceanic CO_2 concentration (~10-25 μM), and its oxygenase activity competes with carboxylation (Badger et al., 1998). To overcome difficulties in acquiring CO_2 from external environment and deficiencies of RubisCO, marine diatoms as well as other microalgae developed CO_2 concentrating mechanisms (CCMs), which actively take up HCO_3^- and CO_2 and elevate the CO_2 concentration around RubisCO. The CCM's capacity is greatly upregulated under low CO_2 conditions (Colman and Rotatore, 1995; Badger et al., 1998; Kaplan and Reinhold, 1999). Biophysical CCMs, which rely solely on the pumping and accumulation of inorganic carbon, are most common but some microalgae have C_4-like biochemical CCMs, which incorporate CO_2 into four carbon organic intermediates (see section1.2.4).

The common features of biophysical CCMs are uptake and accumulation of DIC within the cell and conversion of HCO_3^- to CO_2 around RubisCO, raising the concentration of CO_2 in the immediate vicinity of the enzyme. The main components of biophysical CCMs include inorganic carbon uptake systems, carbonic anhydrases (enzymes that catalyze the reversible

dehydration of HCO_3^- to CO_2) , as well as a microcompartment where RusbiCO aggregates. C_4 CCMs in microorganisms have additional machinery to generate a four carbon organic intermediate (a C_4 compound) from a C_3 compound and a molecule of C_i (inorganic carbon, either CO_2 or HCO_3^-), a way to transport the C_4 compound into the chloroplast, and a mechanism to regenerate C_i from the C_4 compound. Considering different CCM components, their locations and activities, microalgae possess a diverse architecture of CCMs. The physiology and molecular biology of CCMs is best studied in model cyanobacteria and in the model green eukaryotic alga, *Chlamydomonas reinhardtii* (Bader and Price, 2003; Price et al, 2008; Moroney and Ynalvez, 2007; Jungnick et al., 2014). These model systems provide some insight into the CCMs of diatoms.

1.2.2 CCMs in Cyanobacteria

Our understanding of the physiology and molecular biology of the CCM in model cyanobacteria is advanced and illustrates important principles of functional, efficient CCMs (Badger et al., 2002). The essential components of cyanobacteria CCMs are active uptake systems for both CO_2 and HCO_3^- and the carboxysome microcompartment, where CO_2 can be generated via carbonic anhydrase in close proximity to RubisCO (Figure 1.1). Cyanobacteria can be divided in two major groups based on RubisCO and carboxysome lineages as α-cyanobacteria (oceanic cyanobacteria), which contain Form 1A RubisCO and α-carboxysomes, and β-cyanobacteria (primarily freshwater species) with Form 1B Rubisco and β-carboxysomes (Tabita, 1999; Badger et al., 2002; Badger and Price, 2003).

HCO_3^- is brought into the cell by active transporters embedded in the cytoplasmic membrane. There are three different families of HCO_3^- transporters in cyanobacteria powered by ATP or sodium gradients (Badger et al., 2006). CO_2 uptake is driven by a concentration gradient

between the external environment and the cell, a gradient that is generated by the active conversion of CO_2 to HCO_3^- in the thylakoid membrane. This system is not fully understood, but it relies on energy from NADPH to drive the reaction (Kaplan and Reinhold, 1999). These import processes work to accumulate HCO_3^- in the cytoplasm, since CA is absent and the pH is relatively high (7.8- 8.2) (Price and Badger, 1989). Accumulated HCO_3^- then diffuses into the carboxysome where RubisCO is localized. In the carboxysome microcompartment, HCO_3^- is equilibrated with CO_2 by CAs within the carboxysome shell. Both types of carboxysomes contain a specific CA to catalyze the conversion of HCO_3^- to CO_2. β-carboxysomes possess CcaA/IcfA, while in α-carboxysomes the shell protein CsoS3 (CsoSCA) is required (Fukuzawa et al., 1992; Yu et al., 1992; So et al., 2002; So et al., 2004). This equilibration raises the concentration of CO_2 around RubisCO helping it to work at its maximal rate and minimizing oxygenation. Additionally the carboxysome helps to prevent diffusive leakage of CO_2 due to its small size and perhaps as a result of some specific adaptations. Analysis of the carboxysomal CA CsoS3 (CsoSCA) activity indicates the presence of a leak barrier associated with the carboxysome shell; the CA may recycle CO_2 leaking from the carboxysome back into HCO_3^- before it leaves the cell (Badger et al., 2006; Price et al., 2008).

1.2.3 CCMs in *Chlamydomonas reinhardtii*

The CCMs of eukaryotes are diverse, and one of the most extensively characterized is that of the model eukaryotic green alga *Chlamydomonas reinhardtii*. The CCMs of *C. reinhardtii* and cyanobacteria have some features in common, such as the accumulation of HCO_3^-, and confinement of RubisCO to a protein microcompartment (the pyrenoid in eukaryotes), but there are some important differences. Multiple HCO_3^- transporters in the plasma membrane (LCI1 and HLA3) and the chloroplast envelope (NAR1.2 and CCP1/2) take up HCO_3^- from the external

environment and deliver it to the chloroplast. CO_2 uptake across the plasma membrane occurs by diffusion, possibly aided by CO_2 channels (RHP1 and RHP2) (Duanmu et al., 2009; Ohnishi et al., 2010; Miura et al., 2004; Ramazanov et al., 1993; Amoroso et al., 1998). Through the action of these transporters and various CAs, HCO_3^- is accumulated in the chloroplast stroma. HCO_3^- is then transported into the thylakoid lumen where the low pH of the compartment (~5) and presence of CA (CAH3) drive the conversion of HCO_3^- to CO_2 (Moroney and Ynalvez, 2007; Merchant et al., 2007; Ynalvez et al., 2008).

Unlike cyanobacteria where the co-localization of a CA that generates CO_2 and RubisCO in the carboxysome is clearly designed for efficient coupling of the reactions, the location of CO_2 generation and consumption are spatially distinct in *C. reinhardtii*. In *C. reinhardtii*, there is no CA in the pyrenoid, the eukaryotic microcompartment analogous to the cyanobacterial carboxysome. Instead CO_2 generated in the thylakoid lumen must diffuse into the pyrenoid for fixation (Karlsson et al., 1998). A portion of the thylakoid lumen penetrates the pyrenoid providing some overlap between CO_2 generation and consumption, but the coupling does not appear to be as tight as in cyanobacteria.

1.2.4 CCMs in diatoms

Like cyanobacteria and green algae, marine diatoms actively take up both CO_2 and HCO_3^- for photosynthesis and they possess CO_2 concentrating mechanisms to overcome the limited availability of CO_2 in seawater (Colman and Rotatore, 1995; Johnston and Raven, 1996; Matsuda et al., 2001; Trimborn et al., 2008). In addition to the more widespread biophysical CCM, some diatom species like *Thalassiosira weissflogii* are thought to have a C_4-like biochemical CCM (Reinfelder et al., 2000; Roberts et al., 2007; McGinn and Morel, 2008). The main feature distinguishing this system from the biophysical CCM is the role of a C_4 organic

compound (typically malate or oxaloacetate) as an intermediate that is accumulated and transported through the cell. Although there are genes encoding C_4 pathway enzymes in the diatoms *T. pseudonana* and *P. tricornutum*, sequence analysis indicates that the decarboxylating enzyme phosphoenol pyruvate carboxykinase (PEPCK) and malic enzyme do not possess plastid targeting sequences, which would be expected in a functional C_4-CCM (Kroth et al., 2008). Although their role is not always understood, these genes are commonly found in photosynthetic organisms, even those that definitely do not have a C_4 photosynthetic pathway. Radiotracer experiments showed little labeling of C_4 compounds in *T. pseudonana* indicating that it is likely to have a biophysical CCM (Roberts et al., 2007). Inactivation of a critical C_4 enzyme (pyruvate-orthophosphate dikinase) using RNAi provided further evidence that *P. tricornutum* uses a biophysical CCM (Haimovich-Dayan et al., 2013). Therefore, the general occurrence of C_4 biochemical CCMs in diatoms is still in question.

Biophysical CCMs are thought to be more common in diatoms and the biophysical CCM of the model marine diatom *Phaeodactylum tricornutum* is shown in Figure 1.2. HCO_3^- is brought into the cell using active, membrane-embedded transporters, whereas CO_2 uptake is driven by a concentration gradient, since the high permeability of biological membranes to CO_2 precludes the use of active transporters. HCO_3^- is accumulated in the cell (in the chloroplast stroma in Figure 1.2), and then converted to CO_2 by CAs in the vicinity of RubisCO where CO_2 is fixed. RubisCO is typically localized to the pyrenoid, a microcompartment composed primarily of semi-crystalline RubisCO, and the small volume of this compartment helps to limit the diffusive efflux of CO_2. Compartmentalization of CA is crucial to the operation of the CCM (Tachibana et al., 2011; Samukawa et al., 2014). CA cannot be present in compartments where HCO_3^- is accumulated since otherwise the accumulated HCO_3^- would be equilibrated with CO_2,

which can diffuse out of the cell. But high CA activity is required in regions where CO_2 is generated (around RubisCO) or recovered (the cytoplasm) since the uncatalyzed rate of CO_2/HCO_3^- interconversion is slow (Hopkinson, 2014).

CCMs in diatoms have been studied most intensively in two model diatoms, *Thalassiosira pseudonana* (a centric diatom) and *Phaeodactylum tricornutum* (a pennate diatom) since they were the first two diatoms for which nuclear genomes were fully sequenced and for which the most advanced genetic manipulation systems are available (Armbrust et al., 2004; Bowler et al., 2008). Candidate HCO_3^- transporters are found in the genomes of both model diatoms. Seven out of ten putative HCO_3^- transporters found in *P. tricornutum* have been shown to be similar to members of the mammalian solute carrier 4 (SLC4) family, and the rest are similar to the solute carrier 26 (SLC26) family (Nakajima et al., 2013). Homologs also exist in *T. pseudonana*, suggesting SLC transporters might be a common feature in marine diatoms. Phylogenetic analysis showed that some SLC4 family HCO_3^- transporters from these two model diatoms likely share a common origin with human SLC4 genes. One of the putative SLC4 family HCO_3^- transporters, PtSLC4-2 has been localized to the plasma membrane of *P. tricornutum* and was demonstrated to directly pump HCO_3^- from seawater. PtSLC4-2 is sodium dependent, and is most likely either a Na^+- HCO_3^- co-transporter or a Na^+ -driven Cl^-/ HCO_3^- exchanger (Nakajima et al., 2013). Other potential HCO_3^- transporters have been predicted to be localized to the chloroplast membrane system based on analysis of signal peptides in both *P. tricornutum* and *T. pseudonana* (Kroth et al., 2008). Though the location and function of these HCO_3^- transporters have not been confirmed yet, they may work to provide a route for HCO_3^- uptake from the external environment and transport into chloroplast.

Another important component of the CCM are carbonic anhydrases. The subcellular location and activity of CAs are critical to efficient operation of CCMs. In *P. tricornutum*, two β-CAs have been confirmed to be localized in the pyrenoid and to be upregulated at low CO_2 (Harada et al., 2005). Localization to the pyrenoid suggests that these two β-CAs play a key role in the CCM, converting HCO_3^- to CO_2 in close proximity to RubisCO (Satoh et al., 2001; Tanaka et al., 2005; Harada and Matsuda, 2005; Tachibana et al., 2011). Besides these two β-CAs, five α-CAs have been localized to the four-layered chloroplastic membrane system. These CAs may serve to balance the DIC flux into and out of the chloroplast, including recovering CO_2 leaking out of the pyrenoid and converting CO_2 diffusing into the cell to HCO_3^-. Given the possible HCO_3^- transporters localized to the chloroplast membrane system, as well as the lack of stromal CA, HCO_3^- is presumably accumulated in the stroma and finally transported to pyrenoid, where it is converted to CO_2 by CA near RubisCO (Tachibana et al., 2011).

Both CCM genetics and physiology are relatively well-understood in *P. tricornutum*, leading to an integrated description of the CCM in *P. tricornutum* (Tachibana et al., 2011; Hopkinson 2014). However, recent work on CA localization in *T. pseudonana* has suggested the CCM may work differently in this species (Samukawa et al., 2014). In contrast to the CA distribution in *P. tricornutum*, there are no CAs identified in *T. pseudonana*'s pyrenoid. Instead a CA is present in the chloroplast stroma, which would be expected to prevent the accumulation of HCO_3^- in this compartment. A γCA was found to be localized in the cytoplasm and it presumably functions to convert CO_2 that enters the cytoplasm either from external environment or from the chloroplast to HCO_3^-. Moreover, it was reported that *T. pseudonana* possessed two extracellular CAs, which are associated with the cell surface and are lacking in *P. tricornutum*. These eCAs are used to convert HCO_3^- to CO_2 at the cell surface to support CO_2 uptake into the cell. The

ubiquity of eCA in centric marine diatoms implies that it is a component of the CCM in this group (Hopkinson et al., 2013).

Overall, the different subtypes, locations, and activities of those CCM genes that have been identified so far suggest a diversity of CCM architecture in the two model diatoms. There are still molecular details and functions of CCM components that remain unknown and further investigation is needed.

1.2.5 CCM evolution

Oxygenic photosynthetic cyanobacteria evolved at least 2.4 Gya and a wide range of eukaryotic lineages have evolved since 1.2-1.1 Gya (Rasmussen et al., 2008; Butterfield, 2000&2004; Butterfiled et al., 1988). It is likely the earliest RubisCOs were operating in a high CO_2 environment, leading to a low CO_2 affinity or low CO_2/O_2 selectivity (Tabita et al., 2007 and 2008; Raven et al., 2012). The majority of oxygenic organisms used RubisCO in the photosynthetic carbon fixation pathway early in their evolutionary history and nearly all still do so today (Hohmann-Marriott and Blankenship, 2011). Early on, algae presumably relied upon diffusive CO_2 entry to supply RubisCO, since CO_2 concentrations were high (Raven et al., 2012).

Later as O_2 built up in the ocean and atmosphere, the oxygenase activity of RubisCO (photorespiration) began to be relevant, revealing one of the problematic characteristics of this enzyme. The product of RubisCO oxygenase activity, besides 3-phosphoglycerate, is 2-phospohglycolate, which requires additional enzymes and energy to metabolize (Bauwe et al., 2010; Raven et al., 2012). At the same time, decreasing CO_2 made it more difficult for algae to depend solely on diffusive CO_2 entry to supply RubisCO. To overcome the inefficiencies of RubisCO and the additional energy cost of photorespiration, CCMs evolved as an important evolutionary response (Raven et al., 2011 and 2012; Beerling, 2012).

There were several times in Earth's history (2.4-2.1Ga; at 0.75, 0.6 Ga; 320-270 Ma and the Pleistocene 2.1 Ma) during which low CO_2 conditions might have yielded selective pressure for CCMs (Giordano et al., 2005; Raven et al., 2011). It is unclear exactly when CCMs first evolved, but CCMs in modern cyanobacteria and eukaryotes almost certainly evolved independently, and most evidence suggests CCMs evolved independently within the major eukaryotic algal lineages (Badger et al., 1998 and 2002; Badger and Price, 2003). No direct fossil evidence exists and molecular clocks have proven unhelpful regarding the origin of CCMs (Giordano et al., 2005; Raven et al., 2008, 2012). Young et al. (2012) reported that there was positive selection on form ID RubisCO in Bacillariophyta (diatoms) and Haptophyta that corresponded to low CO_2 episodes in the geological record, which possibly corresponds to the time of emergence of CCMs in these lineages.

If some algal CCMs emerged before the Pleistocene (2.1 Ma), then the question arises as to why and how CCMs have survived during intervening periods of high temperature and high CO_2. Raven et al. (2011 and 2012) pointed out that despite increased CO_2 concentrations in the future, changes in other environmental factors will favor retention of CCMs. Under a high CO_2 and high temperature environment, which might be expected in the future, DIC and H^+ concentration will increase in surface seawater, while the depth of upper mixed layer decreases. Taken together, the availability of DIC, PAR (photosynthetically active radiation), and UV-B (Ultraviolet B radiation) increases while the supply of nutrients like N and P decreases in surface water. Short-term experiments on extant phytoplankton revealed that the decreased nutrient supply, with the increased mean flux of PAR and UV-B will lead to upregulation of CCMs even under high CO_2 conditions (Giordano et al., 2005; Raven et al., 2005b, 2008, 2011&2012; Beardall et al., 2009a and 2009b). Much work is still needed on the regulation of CCMs in

response to changes in environmental factors other than CO_2. Exploring CCM evolution can shed light on how CCMs were retained in the past high CO_2 period and how present day CCMs will response to such environments that are expected in the future.

1.3 Objectives and Chapter Overview

The main focus of this dissertation is to understand CO_2 concentrating mechanisms (CCMs) in marine diatoms from genetic, physiological, and evolutionary perspectives. This dissertation work employs genetic techniques, physiological assessment, and bioinformatic approaches. In the chapters that follow, I aim to characterize CCM components in two model diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, from localization to function, and to explore the diversity of CCM components among other diatoms.

The second chapter is focused on characterizing carbonic anhydrases and HCO_3^- transporters in *T. pseudonana* and *P. tricornutum* from subcellular localization to physiological assessment. Understanding CCMs in model diatoms requires knowing where the CCM components are localized. With gene transformation techniques and fluorescence microscopy, I localized a putative CA and three HCO_3^- transporters in *T. pseudonana*, which helped predict their roles in the CCM based on their sub-cellular localization. Mutant lines overexpressing selected CCM components were characterized for function using physiological assays involving Membrane Inlet Mass Spectrometry (MIMS). Using methods such as ^{18}O exchange (Tu et al., 1978; Hopkinson et al., 2011) or photosynthesis vs. inorganic carbon (Ci) concentration (Badger et al., 1994), CA activity, photosynthetic rates, and HCO_3^- uptake were assessed and compared with those of wild type lines to demonstrate the activity and function of the target protein. The genetic and physiological characterization of CA and HCO_3^- transporter are expected to help build a more complete picture of CCMs in *T. pseudonana* and *P. tricornutum*.

The third chapter focuses on exploring the function of extracellular carbonic anhydrase (eCA), an important CCM component that serves to convert HCO_3^- to CO_2 at the cell surface. eCA is unusually common in centric diatoms and I investigated eCA activity in six centric diatom species spanning nearly the full range of cell sizes for centric diatoms (equivalent spherical radius 3-67 μm) using MIMS to assesses eCA's importance for photosynthetic CO_2 supply. Since larger cells are more susceptible to diffusion limitation, I hypothesized 1) that eCA activity would increase with cell size to support greater CO_2 demand by larger cells and 2) that eCA activity would be of greater importance to inorganic carbon supply in larger cells. eCA activity was quantified using a method we recently developed, which allows direct comparison of diffusive CO_2 fluxes with CO_2 generation rates from eCA (Hopkinson et al., 2013). The effect of eCA activity on photosynthesis was assessed by comparing the carbon fixation rates with and without an eCA inhibitor. This work used quantitative measurements to demonstrate the importance of eCA to photosynthesis in larger centric diatoms.

The CCM has been best studied in two model diatoms, and the characterization of CAs and HCO_3^- transporters indicates a diversity of CCM architecture between these two model species (Tachibana et al., 2011; Samukawa et al., 2014; Nakajima et al., 2013). The purpose of Chapter 4 is to explore how similar or different CCM components are among diatoms based on genome and transcriptome data from 34 diatom strains, representing a wide variety of marine diatoms. An additional goal of the work is to determine if the CCMs of the two model diatoms are representative of diatoms in general. I used bioinformatic approaches to screen possible CCM genes in diatom genomes and transcriptomes, and then clustered sequences of each CCM component into groups by OthoMCL and phylogenetic methods. The CCM gene content of these diatoms was compared using hierarchical clustering. Finally, the similarity of CCM gene content

(as represented by the hierarchical clustering) was compared to the phylogenetic relationships among the species (based on a 18S rDNA phylogenetic tree) to determine if CCM diversity tracked phylogeny.

References

Amoroso, G., Saltemeyer, D., Thyssen, C., Fock, H. P. 1998. Uptake of HCO_3^- and CO_2 in cells and chloroplasts from the Microalgae *Chlamydomonas reinhardtii* and *Dunaliella tertiolecta*. *Plant Physiol*. 116(1): 193-201.

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., Zhou, S., Allen, A. E., Apt, K. E., Bechner, M., Brzezinski, M. A., Chaal, B. K., Chiovitti, A., Davis, A. K., Demarest, M. S., Detter, J. C., Glavina, T., Goodstein, D., Hadi, M. Z., Hellsten, U., Hildebrand, M., Jenkins, B. D., Jurka, J., Kapitonov, V. V., Kröger, N., Lau, W. W., Lane, T. W., Larimer, F. W., Lippmeier, J. C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M. S., Palenik, B., Pazour, G. J., Richardson, P. M., Rynearson, T. A., Saito, M. A., Schwartz, D. C., Thamatrakoln, K., Valentin, K., Vardi, A., Wilkerson, F. P., Rokhsar, D. S. 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*. 306: 79-86.

Badger, M. R., Palmqvist, K., Yu, J-W. 1994. Measurement of CO_2 and HCO_3^- fluxes in cyanobacteria and microalgae during steady-state photosynthesis. *Physiol Plant*. 90: 529-536.

Badger, M. R., Andrews, T. J., Whitney, S. M., Ludwig, M., Yellowlees, D. C., Leggat, W., Price, G. D. 1998. The diversity and coevolution of rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. *Canadian Journal of Botany*. 76: 1052-1071.

Badger, M. R., Hanson, D., Price, G. D. 2002. Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. *Funct Plant Biol*. 29: 161-173.

Badger, M. R., Price, G. D. 2003. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot*. 54: 609-622.

Badger, M. R., Price, G. D., Long, B., Woodger, F. 2006. The environmental plasticity and ecological genomics of the cyanobacterial CO_2 concentrating mechanism. *J. Exp. Bot*. 57: 249-265.

Bauwe, H., Hagemann, M. and Fernie, A. R. 2010. Photorespiration: players, partners and origin. *Trends Plant Sci*. 15: 330-336.

Beardall, J., Sobrino, S., Stojkovic, S. 2009a. Interactions between impacts of ultraviolet radiation, elevated CO2, and nutrient limitation in marine primary producers. *Photochem Photobiol Sci*. 8: 1257-1265.

Beardall, J., Stojkovic, S., Larson, S. 2009b. Living in a high CO_2 world; impacts of global climate change on marine phytoplankton. *Plant Ecol Divers*. 2: 191-205.

Beerling, D. J. 2012. Atmospheric carbon dioxide: a driver of photosynthetic eukaryote evolution for over a billion years? *Phil. Trans. R. Soc. B*. 367: 477-482.

Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R. P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J. A., Brownlee, C., Cadoret, J. P., Chiovitti, A., Choi, C. J., Coesel, S., De Martino, A., Detter, J. C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman, M. J., Jenkins, B. D., Jiroutova, K., Jorgensen, R. E., Joubert, Y., Kaplan, A., Kröger, N., Kroth, P. G., La Roche, J., Lindquist, E., Lommer, M., MartinJezequel, V., Lopez, P. J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L. K., Montsant, A., Oudot-Le Secq, M. P., Napoli, C., Obornik, M., Parker, M. S., Petit, J. L., Porcel, B. M., Poulsen, N., Robison, M., Rychlewski, L., Rynearson, T. A., Schmutz, J., Shapiro, H., Siaut, M., Stanley, M., Sussman, M. R., Taylor, A. R., Vardi, A., von Dassow, P., Vyverman, W., Willis, A., Wyrwicz, L. S., Rokhsar, D. S., Weissenbach, J., Armbrust, E. V., Green, B. R., Van de Peer, Y., Grigoriev, I. V. 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*. 456: 239-244.

Broecker, W. S. 1975. Climatic Change- Are we on brink of a pronounced global warming? *Science*. 189: 460- 463.

Butterfield, N. J. 2000. Bangiomorpha n. gen., n. sp.: implications for the evolution of sex, unicellularity, and the Mesoproterozoic/Neoproterozic radiation of eukaryotes. *Palaeobiology*. 26: 386-404.

Butterfield, N. J. 2004. A vaucheriacean alga from the middle Neoproterozoic of Spitsbergen: implication for the evolution of Proterozoic eukaryotes and the Cambrian explosion. *Palaeobiology*. 30: 231-252.

Butterfield, N. J., Knoll, A. H. and Swett, K. 1988. Exceptional preservation of fossils in an upper Proterozoic shale. *Nature*. 334: 424-427.

Caldeira, K. and Wickett, M. E. 2003. Anthropogenic carbon and ocean pH. *Nature*. 425: 365.

Chen, X., Qiu, C. E., Shao, J. Z. 2006. Evidence for K^+-dependent HCO_3^- utilization in the marine diatom *Phaeodactylum tricornutum*. *Plant Physiol*. 141: 731-736.

Colman, B., Rotatore, C. 1995. Photosynthetic inorganic carbon uptake and accumulation in two marine diatoms. *Plant Cell Env*. 18: 919-924.

Doney, S. C. 2006. Oceanography: plankton in a warmer world. *Nature*. 444: 695- 696.

Duanmu, D., Miller, A. R., Horken, K. M., Weeks, D. P., Spalding, M. H. 2009b. Knockdown of a limiting-CO_2-inducible gene HLA3 decreases bicarbonate transport and photosynthetic Ci-affinity in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci*. 106(14): 5990-5995.

Falkowski, P., Scholes, R. J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Hoegberg, P., Linder, S., Mackenzie, F. T., Moore III, B., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., Steffen, W. 2000. The global carbon cycle: a test of our knowledge of Earth as a system. *Science*. 290: 291-296.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., Falkowski, P. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. 281: 237-240.

Finkel, Z. V., Beardall, J., Flynn, K. J., Quigg, A., Rees, T. A. V., Raven, J. A. 2010. Phytoplankton in a changing world: cell size and elemental stoichiometry. *J Plankton Res*. 32: 119-137.

Flynn, K. J., Blackford, J. C., Baird, M. E., Raven, J. A., Clark, D. R., Beardall, J., Brownlee, C., Fabian, H., Wheeler, G. L. 2012. Changes in pH at the exterior surface of plankton with ocean acidification. *Nat. Clim. Change*. 2: 510-513.

Fukuzawa, H., Suzuki, E., Komukai, Y., Miyachi, S. 1992. A gene homologous to chloroplast carbonic anhydrase (icfA) is essential to photosynthetic carbon dioxide fixation by *Synechococcus* PCC7942. *Proc. Natl. Acad. Sci. U. S. A.* 89: 4437-4441.

Gibbs, S. P. 1981. The chloroplast endoplasmic reticulum: structure, function and evolutionary significance. *Int Rev Cytol*. 72: 49-99.

Giordano, M., Beardall, J., Raven, J. A. 2005. CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu Rev Plant Biol.* 56: 99-131.

Goyet, C., Poisson, A. 1989. New determination of carbonic acid dissociation constants in seawater as a function of temperature and salinity. *Deep-Sea Res*. 36: 1635-1654.

Haimovich-Dayan, M., Garfinkel, N., Ewe, D., Marcus, Y., Gruber, A., Wagner, H., Kroth, P. G. and Kaplan, A. 2013. The role of C4 metabolism in the marine diatom *Phaeodactylum tricornutum*. *New Phytologist*. 197: 177-185.

Harada, H., Matsuda, Y. 2005. Identification and characterization of a new carbonic anhydrase in the marine diatom *Phaeodactylum tricornutum*. *Can. J. Bot*. 83: 909-916.

Harada, H., Nakatsuma, D., Ishida, M and Matsuda, Y. 2005. Regulation of the Expression of Intracellular β-Carbonic Anhydrase in Response to CO_2 and Light in the Marine Diatom *Phaeodactylum tricornutum*. *Plant Physiol*. 139(2): 1041-1050.

Hohmann-Marriott, M. F. and Blankenship, R. E. 2011. Evolution of photosynthesis. *Annu. Rev. Plant Biol*. 441: 940-941.

Hopkinson, B. M., Dupont, C. L., Aleen, A. E., Morel, F.M.M. 2011. Efficiency of the CO_2-concentrating mechanism of diatoms. *Proc. Nat. Acad. Sci*. 108: 3830-3837.

Hopkinson, B. M., Meile, C., and Shen, C. 2013. Quantification of extracellular carbonic anhydrase activity in two marine diatoms and investigation of its role. *Plant Physiol*. 162: 1142-1152.

Hopkinson, B. M. 2014. A chloroplast pump model for the CO2 concentrating mechanism in the diatom *Phaeodactylum tricornutum*. *Photosyn Res*. 121(2-3): 223-233.

Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., Johnson, C. A. 2001. IPCC climate change: the scientific basis. Cambridge University Press: Cambridge, UK.

Johnston, A. M., Raven, J. A. 1996. Inorganic carbon accumulation by the marine diatom *Phaeodactylum tricornutum*. *Eur J Phycol*. 31: 285-290.

Jungnick, N., Ma, Y., Mukherjee, B., Cronan, J. C., Speed, D. J., Laborde, S. M., Longstreth, D. J., Moroney, J. V. 2014. The carbon concentrating mechanism in *Chlamydomonas reinhardtii*: finding the missing pieces. *Photosynth Res*. 121: 159-173.

Kaplan, A. and Reinhold, L. 1999. CO_2 concentrating mechanisms in photosynthetic microorganisms. *Annu Rev Plant Physiol Plant Mol Biol*. 50: 539-570.

Karlsson, J., Clarke, A. K., Chen, Z. Y., Hugghins, S. Y., Park, Y. I., Husic, H. D., Moroney, J. V., Samuelsson, G. 1998. A novel α-type carbonic anhydrase associated with the thylakoid membrane in *Chlamydomonas reinhardtii* is required for growth at ambient CO_2. *EMBO J*. 17: 1208-1216.

Kranz, S. A., Eichner, M., Rost, B. 2011. Interactions between CCM and N_2 fixation in *Trichodesmium*. *Photosynth Res*. 109: 73-84.

Kroth, P. G. 2002. Protein transport into secondary plastids and the evolution of primary and secondary plastids. *Int Rev Cytol*. 221: 191–255.

Kroth, P. G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M. S., Stanley, M. S., Kaplan, A., Caron, L., Weber, T., Maheswari, U., Armbrust, E. V., Bowler, C. 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS One*. 3: e1426.

Leggett, J., Pepper, W. J., Swart, R. J., Edmonds, J., Meira Filho, L. G., Mintzer, I., Wang, M. X. and Watson, J. 1992. Emissions Scenarios for the IPCC: an Update. Climate Change. The Supplementary Report to The IPCC Scientific Assessment, Cambridge University Press, UK. 68-95

Mann, D. G. 1999. The species concept in diatoms. *Phycologia*. 38: 437-495.

Matsuda, Y., Hara, T., Colman, B. 2001. Regulation of the induction of bicarbonate uptake by dissolved CO_2 in the marine diatom *Phaeodactylum tricornutum*. *Plant Cell Environ*. 24: 611-620.

Matsuda, Y., Nakajima, K., Tachibana, M. 2011. Recent progresses on the genetic basis of the regulation of CO_2 acquisition systems in response to CO_2 concentration. *Photosynth Res*. 109: 191-203.

Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Marachal-Drouard, L. 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*. 318(5848): 245-250.

McGinn, P. J. and Morel, F. M. M. 2008b. Expression and regulation of carbonic anhydrases in the marine diatom *Thalassiosira pseudonana* and in natural phytoplankton assemblages from Great Bay, New Jersey. *Physiol Plant*. 133: 78-9.

Millero, F. J., Woosley, R., Ditrolio, B., Waters, J. 2009. Effect of ocean acidification on the speciation of metals in seawater. *Oceanography*. 22: 72-85.

Miura, K., Yamano, T., Yoshioka, S., Kohinata, T., Inoue, Y., Taniguchi, F., Asamizu, E., Nakamura, Y., Tabata, S., Yamato, K. T., Ohyama, K., Fukuzawa, H. 2004. Expression profiling-based identification of CO_2- responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Physiol*. 135: 1595-1607.

Moroney, J. V. and Ynalvez, R. A. 2007. Proposed carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*. *Eukaryotic Cell*. 6: 1251-1259.

Moustafa, A., Beszteri, B., Maier, U. G., Bowler, C., Valentin, K., Bhattacharya, D. 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science*. 324: 1724-1726.

Nakajima, K., Tanaka, A., Matsuda, Y. 2013. SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc. Natl. Acad. Sci. U. S. A.* 110: 1767-1772.

Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A. & Quéguiner, B. 1995. Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biochemical Cycles*. 9: 359-372.

Ohnishi, N., Mukherjee, B., Tsujikawa, T., Yanase, M., Nakano, H., Moroney, J. V., Fukuzawa, H. 2010. Expression of a low CO_2- inducible protein, LCI1, increases inorganic carbon uptake in the green alga *Chlamydomonas reinhardtii*. *Plant Cell*. 22: 3105-3117.

Price, G. D., Badger, M. R. 1989. Expression of human carbonic anhydrase in the Cyanobacterium *Synechococcus* PCC7942 creates a high CO_2-requiring phenotype:

Evidence for a central role for carbonxysomes in the CO_2 concentrating mechanism. *Plant physiol*. 91: 505-513.

Price, G. D., Badger, M. R., Woodger, F. J., Long, B. M. 2008. Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): Functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. *J. Exp. Bot*. 59(7): 1441-1461.

Ramazanov, Z., Mason, C. B., Geraghty, A. M., Spalding, M. H., Moroney, J. V. 1993. The low CO_2-inducible 36-kilodalton protein is localized to the chloroplast envelope of *Chlamydomonas reinhardtii*. *Plant Physiol*. 101(4): 1195-1199.

Rasmussen, B., Fletcher, I. R., Brocks, J. J. & Kilburn, M. R. 2008. Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature*. 455: 1101-1104.

Raven, J. A. 1994. Carbon fixation and carbon availability in marine phytoplankton. *Photosynth Res*. 39: 259-273.

Raven, J. A., Brown, K., Mackay, M., Beardall, J., Giordano, M., Granum, E., Leegood, R. C., Kilminster, K. and Walker, D. I. 2005b. Iron, nitrogen, phosphorus and zinc cycling and consequences for primary productivity in the oceans, In Society for General Microbiology Symposium 65 Micro-organisms and Earth systems: Advances in Geobiology (eds. G M Gadd, K T Semple and H M Lappin-Scott). Cambridge University Press, Cambridge. 247-272.

Raven, J. A., Cockell, C. S., La Rocha, C. L. 2008. The evolution of inorganic carbon concentrating mechanisms in photosynthesis. *Phil Trans Roy Soc B.* 363: 2641-2650.

Raven, J. A., Beardall, J., Giordano, M. and Maberly, S. C. 2011. Algal and aquatic plant carbon concentrating mechanisms in relation to environmental change. *Photosynth Res*. 109: 281-296.

Raven, J. A., Giordano, M., Beardall, J. and Maberly, S. C. 2012. Algal evolution in relation to atmospheric CO_2: carboxylases, carbon-concentrating mechanisms and carbon oxidation cycles. *Phil. Trans. R. Soc. B.* 367: 493-507.

Reinfelder, J. R., Kraepiel, A. M. L., Morel, F. M. M. 2000. Unicellular C_4 photosynthesis in a marine diatom. *Nature*. 407: 996-999.

Reinfelder, J. R. 2011. Carbon concentrating mechanisms in eukaryotic marine phytoplankton. *Ann Rev Mar Sci*. 3: 291-315.

Riebesell, U., Schulz, K. G., Bellerby, R. G. J., Botros, M., Fritsche, .P, Meyerhöfer, M., Neill, C., Nondol, G., Oschlies, A., Wohlers, J., Zöllner, E. 2007. Enhanced biological carbon consumption in a high CO_2 ocean. *Nature*. 450: 545-548.

Roberts, K., Granum, E., Leegood, R. C., Raven, J. A. 2007b. C_3 and C_4 pathways of phototsynthetic carbon assimilation in marine diatoms are under genetic, not environmental control. *Plant physiol*. 145: 230-235.

Rost, B., Zondervan, I., Wolf-Gladrow, D. 2008. Sensitivity of the phytoplankton to future changes in ocean carbonate chemistry: current knowledge, contradictions and research directions. *Mar. Ecol. Prog. Ser.* 373: 227-237.

Round, F. E., Crawford, R. M. and Mann, D. G. 1990. The Diatoms: Biology & Morphology of the Genera. Cambridge, UK: Cambridge University Press.

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T. H., Kozyr, A., Ono, T., Rios, A. F. 2004. The oceanic sink for anthropogenic CO_2. *Science*. 305: 367-371.

Samukawa, M., Shen, C., Hopkinson, B. M., Matsuda, Y. 2014. Localization of putative carbonic anhydrases in the marine diatom, *Thalassiosira pseudonana*. *Photosynth Resh*. 121: 235-249.

Satoh, D., Hiraoka, Y., Colman, B., Matsuda, Y. 2001. Physiological and molecular biological characterization of intracellular carbonic anhydrase from the marine diatom *Phaeodactylum tricornutum. Plant Physiol*. 126: 1459.

So, A. K., Cot, S. S. W., Espie, G. S. 2002. Characterization of the C-terminal extension of carboxysomal carbonic anhydrase from *Synechocystis sp* PCC6803. *Funct Plant Biol*. 29: 183-194.

So, A. K. C., Espie, G. S., Williams, E. B., Shively, J. M., Heinhorst, S., Cannon, G. C. 2004. A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell. *J Bacteriol*. 186: 623-630.

Tabita, F. R. 1999. Microbial ribulose 1, 5-bisphosphate carboxylase/oxygenase: a different perspective (Review). *Photosynth Res*. 60: 1-28.

Tabita, F. R., Hanson, T. E., Li, H., Satagopan, S., Singh, J. and Chan, S. 2007. Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.* 71: 576.

Tabita, F. R., Satagopan, S., Hanson,T. E., Kreel, N. E. and Scott, S. S. 2008. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.* 59: 1515-1524.

Tachibana, M., Allen, A. E., Kikutani, S., Endo, Y., Bowler, C., Matsuda, Y. 2011. Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Photosynth Res*. 109: 205-221.

Tanaka, Y., Nakatsuma, D., Harada, H., Ishida, M., Matsuda, Y. 2005. Localization of soluble β-carbonic anhydrase in the marine diatom *Phaeodactylum tricornutum*. Sorting to the chloroplast and cluster formation on the girdle lamellae. *Plant Physiol*. 138: 207-217.

Tréguer, P., Nelson, D. M., Bennekom, A. J., DeMaster, D. J., Leynaert, A., Quéquiner, B. 1995. The silica balance in the world ocean: a reestimate. *Science*. 268: 375-379.

Trimborn, S., Lundholm, N., Thomas, S., Richter, K. U., Krock, B., Hansen, P. J., Rost, B. 2008. Inorganic carbon acquisition in potentially toxic and non-toxic diatoms: the effect of pH-induced changes in seawater carbonate chemistry. *Physiol Plant*. 133: 92-105.

Tu, C., Wynns, G. C., McMurray, R. E., and Silverman, D. N. 1978. CO_2 kinetics in red cell suspensions measured by ^{18}O exchange. *J. Biol. Chem.* 253: 8178-8184.

Volk, T. and Hoffert, M. I. 1985. Ocean carbon pumps: analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2 changes. The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present (Eds Sunquist, E. T. & Broecker, W. S.) Monograph 32: 99-110. (Am. Geophys. Union, Washington DC, 1985).

Wolf-Gladrow, D. A., Riebesell, U., Burkhardt, S., Bijma, J. 1999. Direct effects of CO_2 concentration on growth and isotopic composition of marine plankton. *Tellus B Chem Phys Meterol*. 51: 461-476.

Ynalvez, R. A., Xiano, Y., Ward, A. S., Cunnusamy, K., Moroney, J. V. 2008. Identification and characterization of two closely related β-carbonic anhydrases from *Chlamydomonas reinhardtii*. *Physiol Plant.* 133: 15-26.

Young, J. N., Rickaby, R. E. M., Kapralov, M. V. and Filatov, D. A. 2012. Adaptive signals in algal Rubisco reveal a history of ancient atmospheric carbon dioxide. *Phil. Trans. R. Soc. B.* 367: 483-492.

Yu, J. W., Price, G. D., Song, L., Badger, M. R. 1992. Isolation of a putative carboxysomal carbonic anhydrase gene from the cyanobacterium *Synechococcus* PCC7942. *Plant Physiol*. 100: 794-800.

Zeebe, R. E., Wolf-Gladrow, D. A. 2001. CO_2 in seawater: equilibrium, kinetics, isotopes. Elsevier Science, Amsterdam.

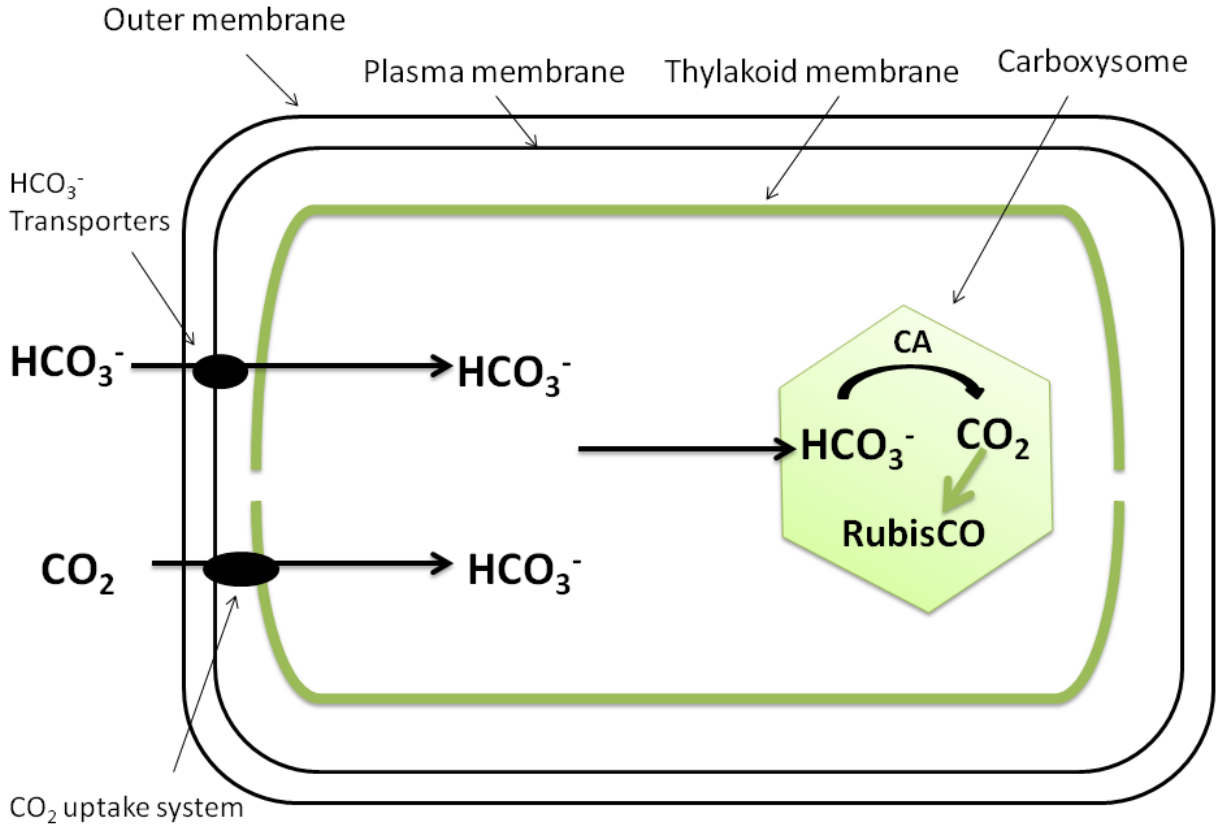


Figure 1.1. A generalized model of the cyanobacterial CCM redrawn according to Price et al. (2008). Green arrow: carbon fiaxation; CA: carbonic anhydrase.

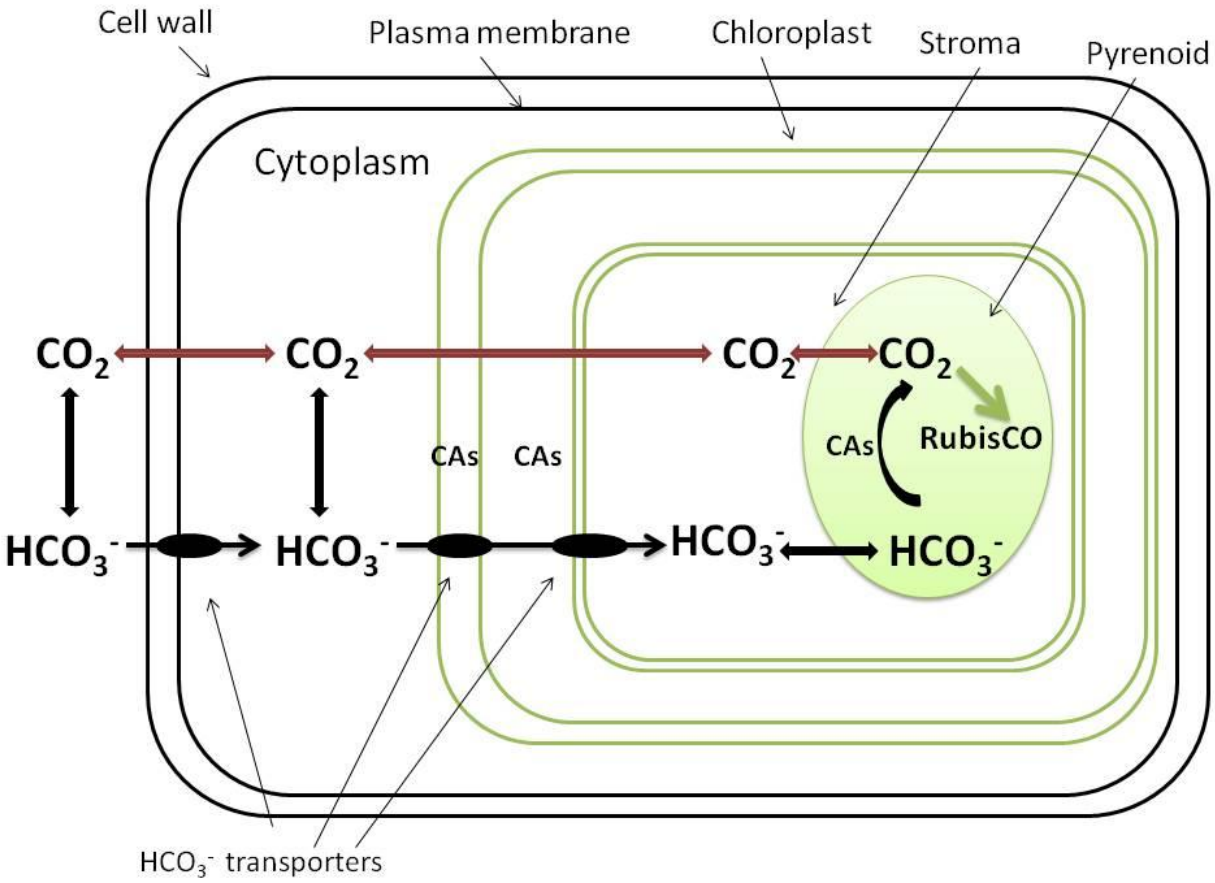


Figure 1.2. A generalized model of the CCM in diatom *Phaeodactylum tricornutum* redrawn according to Hopkinson (2014). Red arrow: diffusive flux; Green arrow: carbon fiaxation; CA: carbonic anhydrase.

CHAPTER 2

GENETIC AND PHYSIOLOGICAL CHARACTERIZATION OF CARBONIC ANHYDRASE AND BICARBONATE TRANSPORTER IN TWO MODEL DIATOMS[1]

[1] Shen, C., Hopkinson, B. M. and Dupont, C. L. To be submitted to *Photosynthesis Research*.

Abstract

Carbonic anhydrases (CAs) and bicarbonate (HCO_3^-) transporters are two crucial components of CO_2 concentrating mechanisms (CCMs) in marine diatoms. Acquisition of HCO_3^- from seawater by HCO_3^- transporters allows algae to overcome CO_2 scarcity in marine environments. HCO_3^- is then accumulated within the cell and converted to CO_2 by intracellular CAs for eventual fixation by RubisCO (Ribulose- 1, 5- bisphosphate carboxylase/oxygenase). Characterization of diatom CCMs is most advanced in the model species *Phaeodactylum tricornutum* and *Thalassiosira pseudonana,* where physiological and genetic work has shed some light on modes of carbon uptake and intracellular transformations. In this study, we fluorescently tagged a homolog of TWCA (TpδCA3), a novel CA family identified in diatoms, in *Thalassiosira pseudonana* to localize the expressed protein. This TWCA homolog formed discrete clusters associated with the chloroplast, most likely in the periplastidal compartment. The activity of this CA was measured in several lines overexpressing the protein, but only one line showed marginally higher internal CA activity. Three putative SLC4 HCO_3^- transporters in *T. pseudonana* were localized via fluorescent protein tagging, one to the plasma membrane and two to the chloroplast. A putative plasma membrane HCO_3^- transporter in *Phaeodactylum tricornutum* was overexpressed and physiologically characterized using membrane inlet mass spectrometry (MIMS) to measure photosynthesis and HCO_3^- uptake as a function of inorganic carbon (Ci) concentration. The diatom lines overexpressing this protein showed higher net photosynthetic rates and HCO_3^- uptake rates compared with wild type at low Ci concentrations indicating that the transporter does indeed transport HCO_3^-. The genetic and physiological characterization of CA and HCO_3^- transporter facilitates building a complete picture of the CCM in *T. pseudonana* and *P. tricornutum*.

1. Introduction

Marine diatoms are responsible for a significant portion of primary production in the ocean and contribute to around 20% of global carbon fixation (Nelson et al., 1995; Field et al., 1998; Mann, 1999). Achieving high rates of photosynthesis requires a plentiful source of inorganic carbon (Ci), and while the total Ci concentration in seawater is ~2 mM, CO_2 concentrations are only ~ 10 μM (Wolf-Gladrow and Riebesell, 1997; Zeebe and Wolf-Gladrow, 2001). This CO_2 concentration is not sufficient to saturate the rate of carbon fixation by Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), the principal enzyme that catalyzes carbon fixation in the Calvin-Benson cycle. While many enzymes work below their maximal rate, keeping RubisCO near saturation is important because it is a very slow enzyme (turnover rates ~5 s^{-1}) but it catalyzes one of the major biochemical reactions in autotrophic cells and even when operating near saturation it uses a significant fraction of the cell's total nitrogen (Losh et al., 2013).

To overcome these difficulties in acquiring CO_2 from seawater, microalgae, including diatoms, have developed systems to actively take up HCO_3^- and CO_2 and elevate the CO_2 concentration around RubisCO (Colman and Rotatore, 1995; Johnston and Raven, 1996; Matsuda et al., 2001; Trimborn et al., 2009). These systems are known as CO_2 concentrating mechanisms (CCMs). Although some species of diatoms may use a C_4 biochemical CCM, most notably *Thalassiosira weisflogii* (Reinfelder et al., 2000), it is generally thought that diatoms use biophysical CCMs, which depend on active pumping of inorganic carbon across cellular membranes (Roberts et al., 2007b). The biophysical CCMs (Figure 2.1) have two major features: accumulation of HCO_3^- within the cell and conversion of that accumulated HCO_3^- to CO_2 around RubisCO. Therefore, the location and activity of HCO_3^- transporters and carbonic anhydrases

(CAs), enzymes that catalyze the reversible dehydration of HCO_3^- to CO_2, are crucial to algal CCMs (Tachibana et al., 2011; Hopkinson et al., 2014; Samukawa et al., 2014).

Identification of components of the biophysical CCM including CAs and HCO_3^- transporters in the model diatoms *Phaeodactylum tricornutum* (a pennate diatom) and *Thalassiosira pseudonana* (a centric diatom) has proceeded rapidly in recent years. CAs are classified into six evolutionarily distinct families: α, β, γ, δ, ε and ζ, of which the δ and ζ families were first identified in diatoms (Roberts et al., 1997; Lane and Morel, 2000). The location and activity of CAs are important for CCM function. At least nine candidate CA genes from the α, β, γ families were identified in the *P. tricornutum* genome (Tachibana et al., 2011). Two β-CAs are localized to the pyrenoid, a protein microcompartment within the chloroplast stroma where RubisCO aggregates (Satoh et al. 2001; Harada and Matsuda 2005; Tachibana et al., 2011). The location and their responsiveness to CO_2 suggest they play an important role in the *P. tricornutum* CCM by converting accumulated HCO_3^- to CO_2 around RubisCO. The presence of pyrenoidal CAs supports the hypothesis that the pyrenoid and carboxysome, which is a protein microcompartment where RubisCO is concentrated in cyanobacteria, may work similarly as a confined compartment to limit the extent of the zone of CO_2 elevation (Price and Badger, 1989). Additional CAs in *P. tricornutum* have been localized to the four-layered chloroplastic membrane system (periplastidal compartment and chloroplastic endoplasmic reticulum) and mitochondria (Tachibana et al., 2011). Five α-CAs, localized to various layers in the chloroplastic membrane system, were shown to be constitutively expressed (Tachibana et al., 2011). These CAs in the chloroplastic membrane system may constitutively control fluxes of Ci. For example, these CAs work to recapture CO_2 leaking out of the chloroplast as HCO_3^-, which can then be transported back into the chloroplast stroma.

Work on the identification of CAs in *T. pseudonana* has been conducted recently. At least thirteen candidate CA genes were found in the genome of *T. pseudonana* (Samukawa et al., 2014). CAs are present in more diverse locations in *T. pseudonana*, including the periplasmic space, mitochondria, chloroplast stroma, and the cytosol (Tachibana et al., 2011; Samukawa et al., 2014). Two CAs were found associated with the cell surface, commonly called extracellular CAs (eCAs), and these enzymes work to convert abundant HCO_3^- in seawater to CO_2 at the cell surface, facilitating CO_2 diffusion into the cells (Hopkinson et al., 2013). A γCA was found in the cytoplasm and its similarity to cyanobacterial γCAs suggests it does function as a CA, unlike some other γCA lineages whose function remains unconfirmed (Pena et al., 2010; Tachibana et al., 2011; Samukawa et al., 2014). This cytosolic CA may serve to convert CO_2 that enters the cytoplasm either from the external environment or from the chloroplast to HCO_3^-, which can then be transported into the chloroplast. No CAs were found in pyrenoid, but one CA was found in the chloroplast stroma, which complicates accumulation of HCO_3^- there and potentially causes greater problems with CO_2 leakage (Samukawa et al., 2014). The actual role of this stromal CA still needs further study.

Putative HCO_3^- transporters, the second key molecular component of CCMs, have been identified in diatom genomes (Kroth et al., 2008; Nakajima et al., 2013). These transporters have significant homology to mammalian solute carrier 4 (SLC4) and solute carrier 26 (SLC26) HCO_3^- transporters. One of these SLC4-type transporters in *P. tricornutum* was localized to the plasma membrane, and was shown to function as a HCO_3^- transporter (Nakajima et al., 2013). Other potential HCO_3^- transporters were predicted to be localized in the chloroplast membrane system based on analysis of signal peptides in both *P. tricornutum* and *T. pseudonana* (Kroth et

al., 2008). These HCO_3^- transporters are thought to provide a route for HCO_3^- uptake into the cytoplasm and from there to the chloroplast for fixation.

Location and activity of CAs and HCO_3^- transporters provide important constraints on CCM function. In this work, one candidate δ-CA in *T. pseudonana* was localized and its activity was assessed with an ^{18}O exchange assay using membrane inlet mass spectrometry (MIMS). Three putative SLC4 HCO_3^- transporters in *T. pseudonana* were localized. Physiological characterization of several putative HCO_3^- transporters in *T. pseudonana* and *P. tricornutum* was attempted using MIMS to measure photosynthesis and HCO_3^- uptake as a function of inorganic carbon concentration.

2. Materials and Methods

2.1 Culture conditions

The diatoms *T. pseudonana* (CCMP1335) and *P. tricornutum* (CCMP 632) were obtained from the National Center for Marine Algae and Microbiota and maintained in the artificial seawater (ASW) medium ESAW (Harrison et al., 1980) modified as described at the Canadian Center for the Culture of Microorganisms website: http://www3.botany.ubc.ca/cccm/NEPCC/esaw.html. The ESAW medium contains 0.55 mM $NaNO_3$ as the sole nitrogen source (nitrate medium). When noted nitrate was replaced by 0.55 mM NH_4Cl, since the CA was placed under the control of the nitrate reductase promoter allowing its expression to be regulated by the availability of nitrate in the medium. EPPS Buffer (5 mM) was added to maintain constant pH, adjusting the initial pH of the culture to be around 7.7 using HCl or freshly prepared NaOH. The pH of the medium was buffered lower than seawater to increase the CO_2 concentration in the medium, in order to suppress expression of native CCM components to allow assessment of overexpressed genes. All cultures were grown at

18°C under fluorescent lights (150 μmol photons· m^{-2} ·s^{-1}) on a 16:8 light: dark cycle. Cell numbers were counted daily using a Coulter Counter, and cells were harvested by filtration during exponential growth.

2.2 Amplifying putative CCM genes in *T. pseudonana*

Carbonic anhydrase gene

The protein sequence of a putative δ-type CA (JGI protein ID 233, named TpδCA3 in a subsequent publication: Samukawa et al., 2014) was aligned using MUSCLE (Edgar, 2004) and shown in Figure 2.2. Reference protein sequences included in the alignment were other putative δ-type CAs from *T. pseudonana* (Samukawa et al., 2014) and the δ-CA from *Thalassiosira weissflogii* TWCA1 (GenBank: AAV39532), the first representative of this family that was discovered (Roberts et al., 1997).

The full length sequence of TpδCA3 was amplified by PCR using genomic DNA as a template. The amplified TpδCA3 sequence was then inserted by blunt end ligation into the pBluescript-TpNR vector. This pBluescript-TpNR vector includes a nitrate reductase promoter, allowing its expression to be regulated by the availability of nitrate in the medium, a green fluorescent protein (GFP) sequence that is joined to the inserted sequence, and a nat1 gene for resistance to the antibiotic nourseothricin used to select diatom transformants containing the vector (Poulsen et al., 2006).

HCO_3^- transporter genes

Sequences of the putative HCO_3^- transporters Tp13887, Tp270240 and Tp23678 (JGI protein ID 13887, 270240, 23678; 13887 is the same as 23678 except that the 5' extension is truncated) were amplified by PCR from cDNA. The amplified sequences were inserted into TpDESTC'YFP vector. This TpDESTC'YFP vector was modified from a Gateway Destination

Vector (Katzen, 2007) to include a fucoxanthin chl a/c binding protein promoter sequence cloned from *T. pseudonana* for constitutive expression, a yellow fluorescent protein (YFP) sequence that is joined to the inserted sequence, and a nat1 gene for resistance to the antibiotic nourseothricin used to select diatom transformants containing the vector.

2.3 Genetic transformation of diatoms

The final vector was introduced into wild-type *T. pseudonana* cells using the Biolistic PDS-1000/He particle delivery system (BIORAD, Hercules, CA, USA) following the protocols of Poulsen et al. (2006). M10 tungsten particles (~3 mg) were coated with 5 μg of circular plasmid DNA using the $CaCl_2$-spermidine method. *T. pseudonana* was grown in ESAW medium and the cells were harvested at a cell density of 10^6 cells mL^{-1}, and further concentrated to 10^8 cells and the cells were plated into a 5 cm diameter circle in the center of an ESAW agar plate (1.5% agar). The cells were then bombarded at a target distance of 7 cm using two rupture disks (1350 psi and 1500 psi). After bombardment, the cells were immediately scraped from the agar plates and resuspended in ESAW liquid medium (one agar plate into 100 mL liquid medium) for recovery overnight with constant illumination. The cells were then concentrated to ~5 x 10^6 cells and plated onto agar plate with the antibiotic nourseothricin (100 μg mL^{-1}). The plates were incubated for 8-14 days with constant light at 18 ℃. Independent colonies were collected from plates and inoculated into ESAW medium supplemented with the same antibiotic and concentration.

2.4 Fluorescence microscopy

A second screening was performed to target cells that showed expression of GFP or YFP by fluorescence microscopy. The resulting GFP/YFP-positive cell lines were used to localize the fluorescently-tagged proteins via confocal microscopy. Confocal imaging of TpδCA3 was

performed using an inverted Zeiss LSM 710 laser scanning microscope. Confocal imaging of Tp13887, Tp270240 and Tp23678 was performed using a Zeiss Axioscope 2 Plus microscope.

2.5 Physiological assessment by Membrane Inlet Mass Spectrometry (MIMS)

To determine if the overexpressed TpδCA3 was active, intracellular carbonic anhydrase (iCA) activity was measured using ^{18}O exchange experiments in the presence of an extracellular carbonic anhydrase (eCA) inhibitor (Hopkinson et al., 2011). In brief, 900 mL ASW/Tris buffer (Ci-free ASW with 20 mM Tris at pH 8.0) was added to the MIMS chamber, followed by 50 μM acetazolamide (AZ), an eCA inhibitor, and 2 mM ^{13}C-^{18}O labeled Ci. The temperature in the chamber was maintained at 20 ℃ using a water jacket. The labeled CO_2 was monitored by MIMS and after ~10 min of monitoring uncatalyzed ^{18}O removal, cells were added from a concentrated suspension. ^{18}O removal by iCA was monitored in the dark for approximately 15 min. The ^{18}O depletion data were then analyzed quantitatively by fitting a model in which iCA activity and the CO_2 and HCO_3^- mass transfer coefficients are unknowns to the data (Tu et al., 1978; Hopkinson et al., 2011). The model treats the interior of the cell as single compartment with iCA activity distributed homogenously throughout the cell. *T. pseudonana* wild type cells and mutant lines (TpδCA3_2, 3, 4, 5) were grown in both nitrate medium and ammonium medium (0.55 mM $NaNO_3$ was replaced with 0.55 mM NH_4Cl) since the pBluescript-TpNR vector is induced by nitrate. ^{18}O exchange experiments were performed on cells grown under these different conditions. The ratio of iCA activity was calculated as the average iCA activity from cells grown in nitrate medium (induced) over that from cells grown in ammonium medium (uninduced). A two-tailed t-test was performed to determine if there was a statistically significant difference between iCA activity in induced vs. uninduced cells.

To assess the function of putative HCO_3^- transporters, net CO_2 uptake, net HCO_3^- uptake and photosynthetic rates were measured as a function of Ci concentration in the medium (Photosynthesis vs. Ci experiment) via a MIMS method (Badger et al., 1994). Briefly, concentrated cells were resuspended in ASW/Tris buffer (pH 8.0) lacking Ci, and were placed in the MIMS chamber. Light (200 µmol photons $\cdot m^{-2} \cdot s^{-1}$) was applied to induce photosynthesis and thus consume any residual Ci. When net photosynthesis stopped (~20 min) the light was turned off and the Photosynthesis vs. Ci experiment was begun. Ci was gradually added to the sample, then photosynthesis and CO_2 and HCO_3^- uptake were measured in a series of light-dark cycles following the method of Badger et al. (1994). Net photosynthetic rates were determined from the rate of O_2 production, and net CO_2 uptake rates were calculated from the extent to which CO_2 was drawn down below equilibrium. Net HCO_3^- uptake was then calculated as the difference between photosynthesis and net CO_2 uptake.

Photosynthesis vs. Ci experiments were performed on *P. tricornutum* wild type and mutant lines (PtBD714, Pt1677) and on *T. pseudonana* wild type and mutant line (Tp23678). A transgenic *P. tricornutum* line containing the BD714 protein expressed under a constitutive promoter was generated by our collaborator, Dr. Chris Dupont at the J. Craig Venter Institute (these transgenic lines are named PtBD714). The protein was localized to the plasma membrane via fluorescent protein tagging. Because BD714 is expressed under a constitutive promoter in the transgenic line it could not be easily turned on and off. Instead, we attempted to suppress the expression of the native HCO_3^- transporters by growing both the wild type and transgenic lines at high CO_2 (pH 7.7) so that the activity of the putative HCO_3^- transporter (BD714) could be assessed. Pt1677 is another transgenic *P. tricornutum* line generated by Dr. Chris Dupont containing the 1677 protein localized to the plasma membrane and the activity of this putative

HCO_3^- transporter was also assessed. Tp23678 is a transgenic *T. pseudonana* line as indicated in section 2.2.

3. Results

Carbonic anhydrase

Thirteen putative CA genes were found in the genome of *T. pseudonana*, four of which are δCAs (Samukawa et al., 2014). Three histidine residues form the Zn binding ligands at the active site of the *Thalassiosira weissflogii* TWCA1 (Cox et al., 2000). Alignment of δCAs homologs (Figure 2.2) showed that some of the potential Zn-binding histidines are well conserved (as also reported by McGinn and Morel, 2008b). Sequence analysis done by Samukawa et al. (2014) showed TpδCA3 lacked any predictable N-terminal presequence, and was thus predicted to be localized to the cytosol. However, the TpδCA3:GFP fusion was found to form several small dots adjacent to the chloroplast autofluorescence (Figure 2.3). The localization of TpδCA3 is most similar to that of other proteins localized to the periplastidal compartment (PPC), a space between the two outer chloroplastic membranes and two inner chloroplastic membranes (Sheiner et al., 2011).

Intracellular carbonic anhydrase activity was measured using ^{18}O exchange experiments on cells grown in both nitrate medium (inducing TpδCA3) and ammonium medium (uninduced control). The ratio of iCA activity in induced vs. uninduced cultures from the wild type and several mutant lines (TpδCA3_2, 3, 4, 5) are shown in Table 2.2. iCA activity of wild type *T. pseudonana* was not significantly affected by the nitrogen source in the growth medium. When TpδCA3:GFP was induced in the mutant lines, one strain had significantly higher iCA acitivity compared with un-induced cells, but in the remaining lines there was no significant increase in iCA activity when TpδCA3 was induced.

HCO_3^- transporters

The functionality of putative HCO_3^- transporters was assessed by comparing photosynthetic rates and HCO_3^- uptake rates in wild type and mutant lines. Net photosynthetic rates and HCO_3^- uptake rates of *P. tricornutum* wild type and PtBD714 mutant lines were obtained from Photosynthesis vs. Ci experiments (Figure 2.4, Figure 2.5, Table 2.3). Both lines overexpressing BD714 showed higher photosynthetic affinity for Ci compared with the wild type, as indicated by the half-saturation constants and initial slope of Michaelis-Menten fits to the data (Table 2.3, Figure 2.4). At low Ci concentrations, HCO_3^- uptake rates in both mutant lines were significantly higher relative to the wild type (Figure 2.5). The same assays were also conducted on another mutant line in which a plasma membrane-localized HCO_3^- transporter was over-expressed (Pt1677, again generated by Dr. Chris Dupont) and on Tp23678, a *T. pseudonana* line in which a putative HCO_3^- transporter localized to the plasma membrane was overexpressed. However, no significant differences in photosynthetic rates or HCO_3^- uptake rates were found between wide type and mutant lines for these proteins (data not shown).

Three putative HCO_3^- transporters were localized using fluorescent protein tagging in *T. pseudonana* (Figure 2.6). Tp23678 was localized to the plasma membrane. The localization of Tp13887, the same gene as Tp23678 except with a truncated 5' extension, was unclear due to low expression of the fluorescent protein. It appears to be localized to the chloroplast, but in any case it is definitely not in the plasma membrane. Tp270240 was localized to the chloroplast membrane system, though exactly which membrane it resides in is unclear.

4. Discussion

In marine diatoms, CA localization and function have been most well studied in the model diatom *P. tricornutum*. Nine candidate CA sequences have been found in the genomes of

P. tricornutum (Tachibana et al., 2011). Two γ-CAs are localized to mitochondria, which may serve to recapture CO_2 produced by respiration although alternative roles for mitochondrial γ-CAs have been proposed in land plants (Soto et al., 2015). All five α-CAs are localized within the four-layered chloroplastic membrane system, but they are not responsive to low CO_2 concentration. These CAs in the membrane system work to control fluxes of Ci into and out of the chloroplast (Tachibana et al., 2011). They are likely involved in the recapture of CO_2 that leaks out of the chloroplast, converting it to HCO_3^- ; the recovered HCO_3^- is then transported into the chloroplast stroma. Most importantly, two βCAs are present within the pyrenoid and are upregulated under low CO_2 conditions (Harada and Matsuda, 2005; Harada et al., 2005). These CAs catalyze the conversion of accumulated HCO_3^- to CO_2, raising the concentration of CO_2 in the immediate vicinity of RubisCO.

Unexpectedly, the CA distribution in *T. pseudonana* is quite different from that of *P. tricornutum*, and suggests that the CCM in *T. pseudonana* works in a fundamentally different way. Both *T. pseudonana* and *P. tricornutum* have CAs in the chloroplast stroma (TpαCA1 and PtβCA1, PtβCA2, using the nomenclature of Samukawa et al., 2014), but in *P. tricornutum* the two βCAs are restricted to the pyrenoid, a logical arrangement that allows HCO_3^- to be accumulated in the stroma and then converted to CO_2 within the pyrenoid. In contrast, TpαCA1 is present throughout the chloroplast stroma, which would be expected to result in the dissipation of a HCO_3^- pool. Nonetheless, *T. pseudonana* has a functional CCM (Trimborn et al., 2009), though the details of its function remain to be worked out.

Although there are several CAs within the four-layered chloroplast membrane of *P. tricornutum*, no CA had yet been localized to the chloroplast membrane system of *T. pseudonana*. In this study, TpδCA3 was found to form distinct punctate aggregations associated

with the chloroplast, most likely within the periplastidal compartment, despite predictions that it would be in the cytosol based on its lack of any predictable N-terminal presequence (Samukawa et al. 2014). This is so far the only CA found within *T. pseudonana* chloroplast membranes, though there are still several CAs that have yet to be localized (Samukawa et al., 2014). Measurement of iCA activity using ^{18}O-exchanged method by MIMS indicated one mutant line had significantly higher iCA activity when TpδCA3 was induced, but other lines showed no effect of induction, and so definitive verification that TpδCA3 is a functional CA awaits further studies. The expression of the TpδCA3 gene has been shown to be upregulated at low CO_2, providing further evidence that the protein is involved in the CCM (Samukawa et al., 2014). These results suggest TpδCA3 may help recover CO_2 leaking out of the chloroplast, which would occur at higher rates when external CO_2 concentrations are low and a large diffusive gradient exists for CO_2 between the site of fixation and the external environment.

In addition to the extensive research on CAs in diatoms, there have also been exciting developments in the identification of HCO_3^- transporters. Putative HCO_3^- transporters found in the model diatoms *P. tricornutum* and *T. pseudonana* have high similarity to the members of the mammalian SLC4 HCO_3^- transporters. Recent work has shown that one of those mammalian type SLC4 HCO_3^- transporters (BD1806, named PtSLC4-2 in Nakajima et al., 2013) was localized to the plasma membrane and does transport HCO_3^-, as predicted based on sequence homology (Nakajima et al., 2013). Moreover, there are other putative SLC4 HCO_3^- transporters that are predicted to be localized to chloroplast membrane system based on the presence of targeting sequences in both *P. tricornutum* and *T. pseudonana* (Kroth et al., 2008). In the present study, physiological characterizations of several putative HCO_3^- transporters from *P. tricornutum* were attempted. We tested transgenic lines expressing two different plasma membrane

transporters (BD714 and 1677), looking for effects of the over-expressed transporters on net photosynthetic rates and HCO_3^- uptake rates. At low Ci concentrations, two independent lines overexpressing BD714 had higher HCO_3^- uptake rates relative to the wild type, and both mutant lines showed significantly higher photosynthetic affinity for Ci compared with wild type (Table 2.3). These results confirm that BD714 is a plasma membrane HCO_3^- transporter. Comparison of lines overexpressing 1677 with wild type provided no evidence for enhanced HCO_3^- uptake or photosynthesis.

The finding that BD714 is a HCO_3^- transporter together with previous reports that PtSLC4-2 is a functional HCO_3^- transporter (Nakajima et al., 2013) provide increased confidence that the mammalian type SLC4 HCO_3^- transporters found in *P. tricornutum* and other diatoms do actually function as HCO_3^- transporters. In mammals, nearly all SLC4 transporters transport HCO_3^- but because of the phylogenetic distance between mammals and diatoms there was uncertainty regarding the substrate specificity of these proteins in diatoms. This family of proteins appears to be the principal HCO_3^- transporters in diatoms, facilitating HCO_3^- uptake from the external environment into cells.

In addition to work on *P. tricornutum*, we localized three putative SLC4 HCO_3^- transporters in *T. pseudonana*, one to the plasma membrane and two to the chloroplast (Figure 2.6). The potential HCO_3^- transport activity of one plasma membrane-localized protein (23678) was assessed, but the net photosynthetic rate and HCO_3^- uptake rate did not show any significant differences compared with the wild type. The activity of transporters localized to the chloroplast membranes has not been confirmed, since it would be challenging to observe the effects of an overexpressed chloroplastic HCO_3^- transporter with the same techniques used to study plasma membrane transporters. When testing the activity of plasma membrane HCO_3^- transporter, the

cells were grown at high CO_2 to suppress the expression of other CCM components, hopefully allowing detect the effects of the overexpressed protein. While this strategy proved useful for a plasma membrane transporter, overexpression of a chloroplastic transporter may not lead to enhanced photosynthesis if the cell is limited by the Ci influx to the cell. Assessment of chloroplast transporters may require heterologous expression, for example in *Xenopus* embryos (Romero et al., 2000), or may require co-expression of a plasma membrane HCO_3^- transporter to observe increased rates of photosynthesis or HCO_3^- uptake.

Although the activity of these HCO_3^- transporters localized to the chloroplast membranes have not been characterized, they likely function as HCO_3^- transporters given functional validation of similar SLC4 transporters in *P. tricornutum*. Assuming this, the SLC4 transporters provide a route for HCO_3^- uptake from the external environment and transport into chloroplast. Since *T. pseudonana* possesses eCA to facilitate CO_2 uptake, it would be interesting to determine if the plasma membrane HCO_3^- transporter plays as important a role as it does in *P. tricornutum*.

Genetic and physiological research on HCO_3^- transporters and CAs has led to a more complete picture of the CCM in model diatoms. In *P. tricornutum*, both CCM genetics and physiology are relatively well-developed and have led to an integrated description of the CCM (Hopkinson, 2014). Here, the discovery that a second SLC4 gene functions as a HCO_3^- transporter provides further confidence that SLC4 homologs are the primary HCO_3^- transporters in diatoms. Identification of a CA within the chloroplast membrane system of *T. pseudonana* (TpδCA3) indicates that this is a common feature among diatoms that is likely involved in the recovery of CO_2 leaking out of the chloroplast. The distinct CCM gene complement and localization in *P. tricornutum* and *T. pseudonana*, in particular the presence of a stromal CA in *T. pseudonana*, suggests that CCMs are diverse among diatoms. Therefore, further investigations of

CAs and HCO_3^- transporters are required, along with complementary physiological work, to develop a comprehensive understanding of the CCMs in diatoms.

References

Badger, M. R., Palmqvist, K., Yu, J-W. 1994. Measurement of CO_2 and HCO_3^- fluxes in cyanobacteria and microalgae during steady-state photosynthesis. *Physiol Plant*. 90: 529-536.

Colman, B., Rotatore, C. 1995. Photosynthetic inorganic carbon uptake and accumulation in two marine diatoms. *Plant Cell Env*. 18: 919-924.

Cox, E. H., McLendon, G. L., Morel, F. M. M., Lane, T. W., Prince, R. C., Pickering, I. J., George, G. N. 2000. The active site structure of *Thalassiosira weissflogii* carbonic anhydrase 1. *Biochemistry*. 39: 12128-12130.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 32(5): 1792-1797.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., Falkowski, P. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. 281: 237-240.

Harada, H., and Matsuda, Y. 2005. Identification and characterization of a new carbonic anhydrase in the marine diatom *Phaeodactylum tricornutum*. *Can. J. Bot*. 83: 909-916.

Harada, H., Nakatsuma, D., Ishida, M and Matsuda, Y. 2005. Regulation of the Expression of Intracellular β-Carbonic Anhydrase in Response to CO_2 and Light in the Marine Diatom *Phaeodactylum tricornutum*. *Plant Physiol*. 139(2): 1041-1050.

Harrison, P. J., Waters, R. E., and Taylor, F. J. R. 1980. A broad spectrum artificial seawater medium for coastal and open ocean phytoplankton. *J. Phycol*. 16: 28-35.

Hopkinson, B. M., Dupont, C. L., Aleen, A. E., Morel, F. M. M. 2011. Efficiency of the CO_2-concentrating mechanism of diatoms. *Proc. Nat. Acad. Sci*. 108: 3830-3837.

Hopkinson, B. M., Meile, C., and Shen, C. 2013. Quantification of extracellular carbonic anhydrase activity in two marine diatoms and investigation of its role. *Plant Physiol*. 162: 1142-52.

Hopkinson, B. M. 2014. A chloroplast pump model for the CO2 concentrating mechanism in the diatom Phaeodactylum tricornutum. *Photosynth Res*. 121: 223-233.

Johnston, A. M., Raven, J. A. 1996. Inorganic carbon accumulation by the marine diatom *Phaeodactylum tricornutum*. *Eur J Phycol*. 31: 285-290.

Katzen, F. 2007. Gateway recombinational cloning: a biological operating system. *Expert Opin Drug Discov*. 4: 571-589.

Kroth, P. G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M. S., Stanley, M. S., Kaplan, A., Caron, L., Weber, T., Maheswari, U., Armbrust, E. V., Bowler, C. 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS One*. 3: e1426.

Lane, T. W., Morel, F. M. M. 2000. A biological function for cadmium in marine diatoms. Proc *Natl Acad Sci USA*. 97: 4627-4631.

Losh, J. L., Young, J. N. and Morel, F. M. M. 2013. Rubisco is a small fraction of total protein in marine phytoplankton. *New Phytologist*. 198: 52-58.

Mann, D. G. 1999. The species concept in diatoms. *Phycologia*. 38: 437-495.

Matsuda, Y., Hara, T., Colman, B. 2001. Regulation of the induction of bicarbonate uptake by dissolved CO_2 in the marine diatom *Phaeodactylum tricornutum*. *Plant Cell Environ*. 24: 611-620.

Matsuda, Y., Nakajima, K., Tachibana, M. 2011. Recent progresses on the genetic basis of the regulation of CO_2 acquisition systems in response to CO_2 concentration. *Photosynth Res*. 109: 191-203.

McGinn, P. J., and Morel, F. M. M. 2008b. Expression and regulation of carbonic anhydrases in the marine diatom *Thalassiosira pseudonana* and in natural phytoplankton assemblages from Great Bay, New Jersey. *Physiol Plant*. 133: 78-79.

Nakajima, K., Tanaka, A., Matsuda, Y. 2013. SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc. Natl. Acad. Sci. U. S. A.* 110: 1767-1772.

Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A. & Quéguiner, B. 1995. Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biochemical Cycles*. 9: 359-372.

Pena, K. L., Castel, S. E., de Araujo, C., Espie, G. S., Kimber, M. S. 2010. Structural basis of the oxidative activation of the carboxysomal γ-carbonic anhydrase, CcmM. *Proc Natl Acad Sci USA*. 107: 2455-2460.

Poulsen, N., Chesley, P. M., Kroger, N. 2006. Molecular genetic manipulation of the diatom *Thalassiosira pseudonana* (Bacillariophyceae). *J.Phycol*. 42: 1059-1065.

Price, G. D., and Badger, M. R. 1989. Expression of human carbonic anhydrase in the Cyanobacterium *Synechococcus* PCC7942 creates a high CO_2-requiring phenotype: Evidence for a central role for carbonxysomes in the CO_2 concentrating mechanism. *Plant physiol*. 91: 505-513.

Reinfelder, J. R., Kraepiel, A. M. L., Morel, F. M. M. 2000. Unicellular C_4 photosynthesis in a marine diatom. *Nature*. 407: 996-999.

Roberts, K., Granum, E., Leegood, R. C., Raven, J. A. 2007b. C_3 and C_4 pathways of photosynthetic carbon assimilation in marine diatoms are under genetic, not environmental control. *Plant physiol*. 145: 230-235.

Roberts, S. B., Lane, T. W., Morel, F. M. M. 1997. Carbonic anhydrase in the marine diatom *Thalassiosira weissflogii* (Bacillariophyceae). *J Phycol*. 33: 845-850.

Romero, M. F., Henry, D., Nelson, S., Harte, P. J., Dillon, A. K. and Sciortino, C. M. 2000. Cloning and Characterization of a Na^+-driven Anion Exchanger (NDAE1). *J. Biol. Chem.* 275: 24552-24559.

Samukawa, M., Shen. C., Hopkinson, B. M., Matsuda, Y. 2014. Localization of putative carbonic anhydrases in the marine diatom, *Thalassiosira pseudonana*. *Photosynthesis Research*. 121: 235-249.

Satoh, D., Hiraoka, Y., Colman, B., Matsuda, Y. 2001. Physiological and molecular biological characterization of intracellular carbonic anhydrase from the marine diatom *Phaeodactylum tricornutum. Plant Physiol*. 126: 1459.

Sheiner, L., Demerly, J. L., Poulsen, N., Beatty, W. L., Lucas, O., Behnke, M. S., White, M. W., Striepen, B. 2011. A systematic screen to discover and analyze apicoplast proteins identifies a conserved and essential protein import factor. *PLoS Pathog*. 7(12): e1002392.

Soto, D., Cordoba, J. P., Villarreal, F., Bartoli, C., Schmitz, J., Maurino, V. G., Braun, H. P., Pagnussat, G. C., Zabaleta, E. 2015. Functional characterization of mutants affected in the carbonic anhydrase domain of the respiratory complex I in *Arabidopsis thaliana*. *Plant Journal*. 83: 831-844.

Tachibana, M., Allen, A. E., Kikutani, S., Endo, Y., Bowler, C., Matsuda, Y. 2011. Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Photosynth Res*. 109: 205–221.

Trimborn, S., Wolf-Gladrow, D., Richter, K-U., Rost, B. 2009. The effect of pCO_2 on carbon acquisition and intracellular assimilation in four marine diatoms. *Journal of Experimental Marine Biology and Ecology.* 376: 26-36.

Tu, C., Wynns, G. C., McMurray, R. E., and Silverman, D. N. 1978. CO_2 kinetics in red cell suspensions measured by ^{18}O exchange. *J. Biol. Chem.* 253: 8178-8184.

Wolf-Gladrow, D. A. and Riebesell, U. 1997. Diffusion and reactions in the vicinity of plankton: a refined model for inorganic carbon transport. Marine Chemistry. 59: 17-34.

Zeebe, R. E., Wolf-Gladrow, D. A. 2001. CO_2 in seawater: equilibrium, kinetics, isotopes. Elsevier Science, Amsterdam

Table 2.1. Primers used for construction of putative CA and HCO_3^- transporters.

Target gene	5’ -oligonucleotides sequence -3’
TpδCA3	Forward primer GTTGATATCATGGGAGACATCACCCCA
	Reverse primer GTTGATATCAGCACGAAGGTTGCGGTT
Tp13887	Forward primer CACCATGAACGGCAAACCAGAAAAGTTC
	Reverse primer TTCTCCCGCTGCAGAAGCAGCCTG
Tp270240	Forward primer CACCATGAGTCTCAAACAACCTCTCAT
	Reverse primer TCCCCATGGCCTTCCACTGGTA
Tp23678	Forward primer CACCATGACCACCTCCGATAATGAC
	Reverse primer TCCTTCTCCCGCTGCAGAAG

Table 2.2. The ratio of iCA activity in *T. pseudonana* wild type and mutant lines with TpδCA3 induced was calculated as the iCA activity of cells grown in nitrate medium over that of cells grown in ammonium medium. A two-tailed t-test was used to test for significant differences between cells in nitrate medium and cells in ammonium medium. The p-value is reported and star (*) means statistically significant different.

	induced/un-induced iCA activity	t-test results
wild type	0.84	0.06
TpδCA3-2	1.30	0.05
TpδCA3-3	1.53	<0.001*
TpδCA3-4	1.25	0.07
TpδCA3-5	1.10	0.27

Table 2.3. Photosynthetic parameters of *P. tricornutum* wild type and Pt BD714, in which a putative plasma membrane HCO_3^- transporter was overexpressed. P_{max} is the maximal photosynthetic rate and $K_{0.5}$ is the half-saturation constant from fits of Michaelis-Menten curves to the data. Initial slope is calculated as $P_{max}/K_{0.5}$. A one-tailed t-test was used to test for significant differences of P_{max} rate between wild type and mutant lines and the p-value is reported. Star (*) means statistically significant different.

	P_{max} ($x10^{-17}$ mol $cell^{-1}$ s^{-1})	$K_{0.5}$ [DIC] (μM)	initial slope	t test
Pt wild type	1.68 ± 0.11	579.2 ± 94.8	0.003	
Pt BD714_9	0.98 ± 0.02	49.6 ± 8.0	0.020	<0.001*
Pt BD714_10	0.86 ± 0.02	67.6 ± 11.6	0.013	0.01*

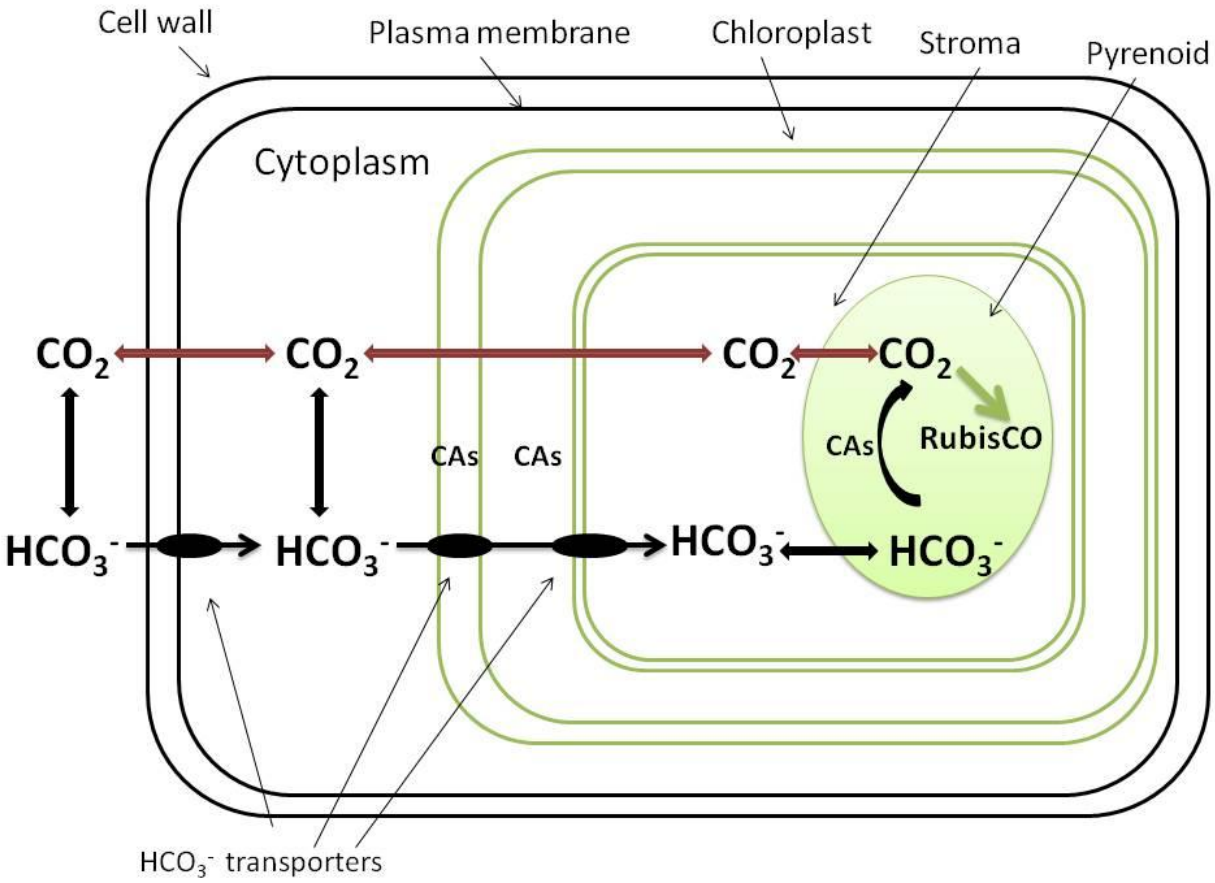


Figure 2.1. A general biophysical CCM presents in *P.tricornutum*. Redrawn according to Hopkinson (2014) and Tachibana et al. (2011). The equilibrium between CO_2 and HCO_3^- is catalyzed by intracellular carbonic anhydrase (CA). Red arrow: diffusive CO_2 fluxes; Green arrow: carbon fixation.

Figure 2.2. Comparison of amino acid sequences of TpδCA3 (named TpδCA_233 in the figure) with putative δ-type CAs from *T. pseudonana* (Samukawa et al., 2014) and δ -type CAs from *Thalassiosira weissflogii* TWCA1 (GenBank: GB acc#, AAV39532). Conserved histidine residues are indicated by star (*).

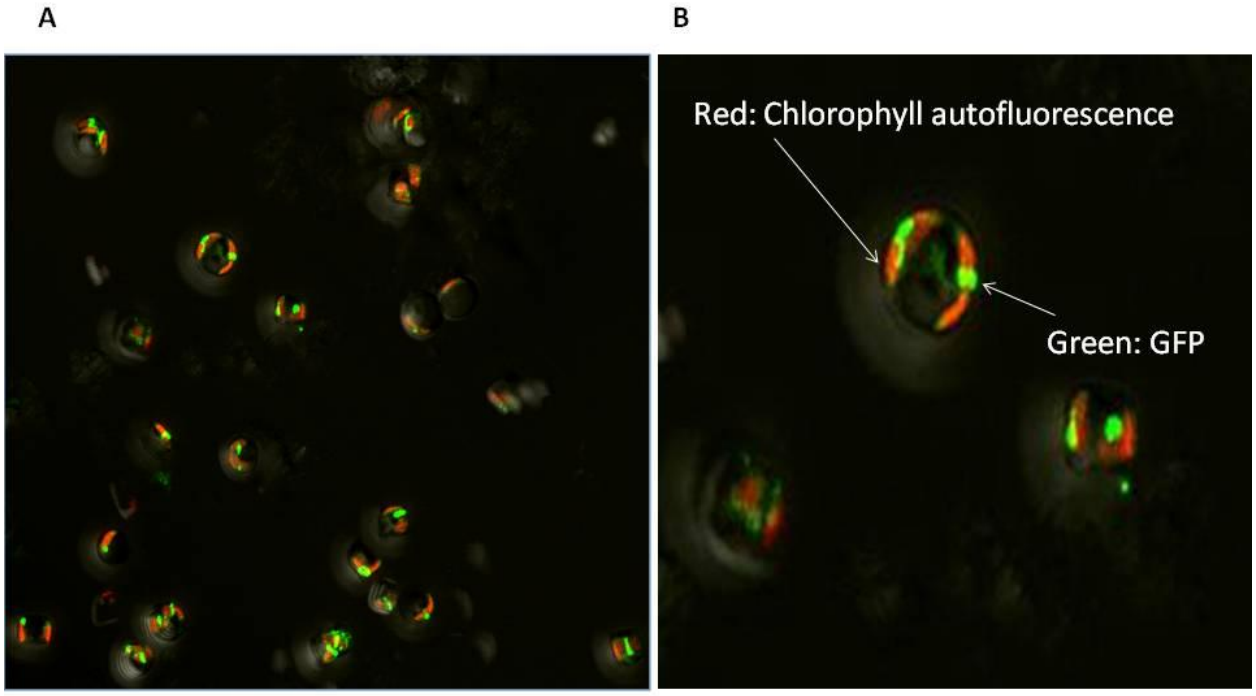


Figure 2.3. Localization of TpδCA3 in *T. pseudonana*. Images were collected on Zeiss LSM 710 Inverted Confocal Microscope. A. Merged channels; Red: Chlorophyll autofluorescence; Green: GFP fluorescence. B. Amplified image from A.

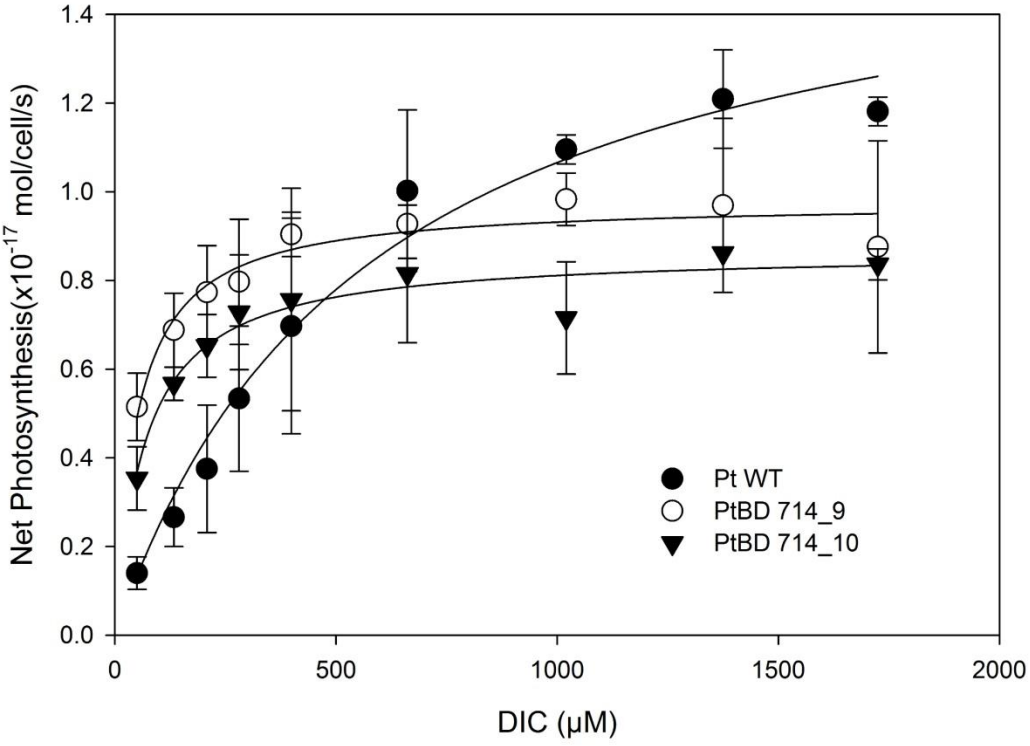


Figure 2.4. Net photosynthesis as a function of DIC (dissolved inorganic carbon, or Ci) concentration in *P. tricornutum* wild type (Pt WT) and mutant lines with plasma membrane HCO_3^-transporter over-expressed (PtBD714_9, PtBD714_10). N = 2 for PtBD714_9 and PtBD71_10 and N = 4 for wild type. The Michaelis-Menten curves are generated to fit the data.

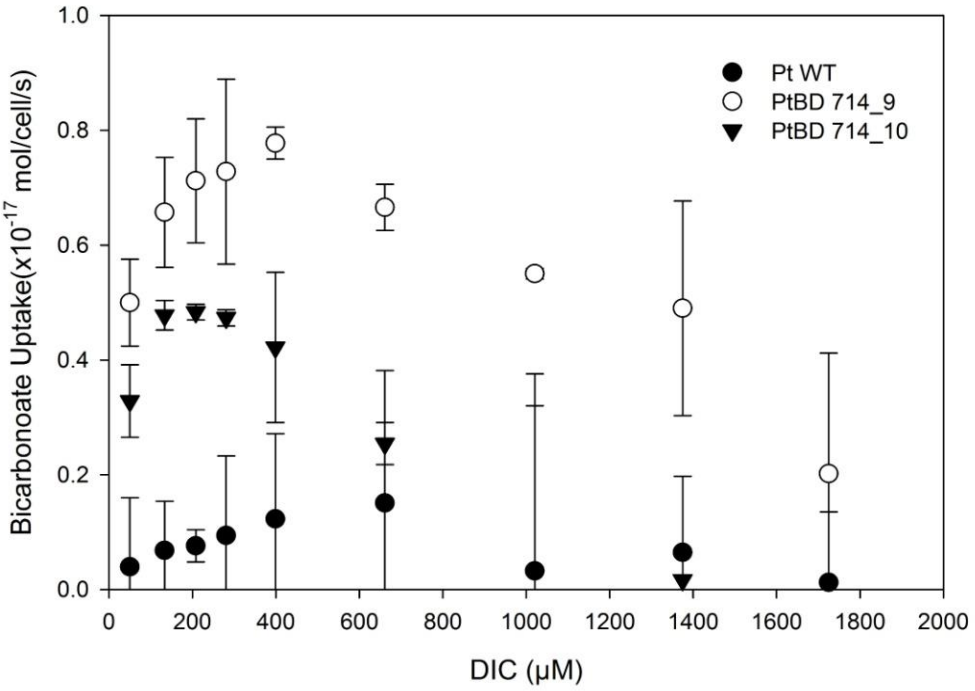


Figure 2.5. Bicarbonate (HCO_3^-) uptake as a function of DIC (dissolved inorganic carbon, or Ci) concentration in *P. tricornutum* wild type (Pt WT) and mutant lines with plasma membrane HCO_3^- transporter over-expressed (PtBD714_9, PtBD714_10). N = 2 for PtBD714_9 and PtBD71_10 and N = 4 for wild type.

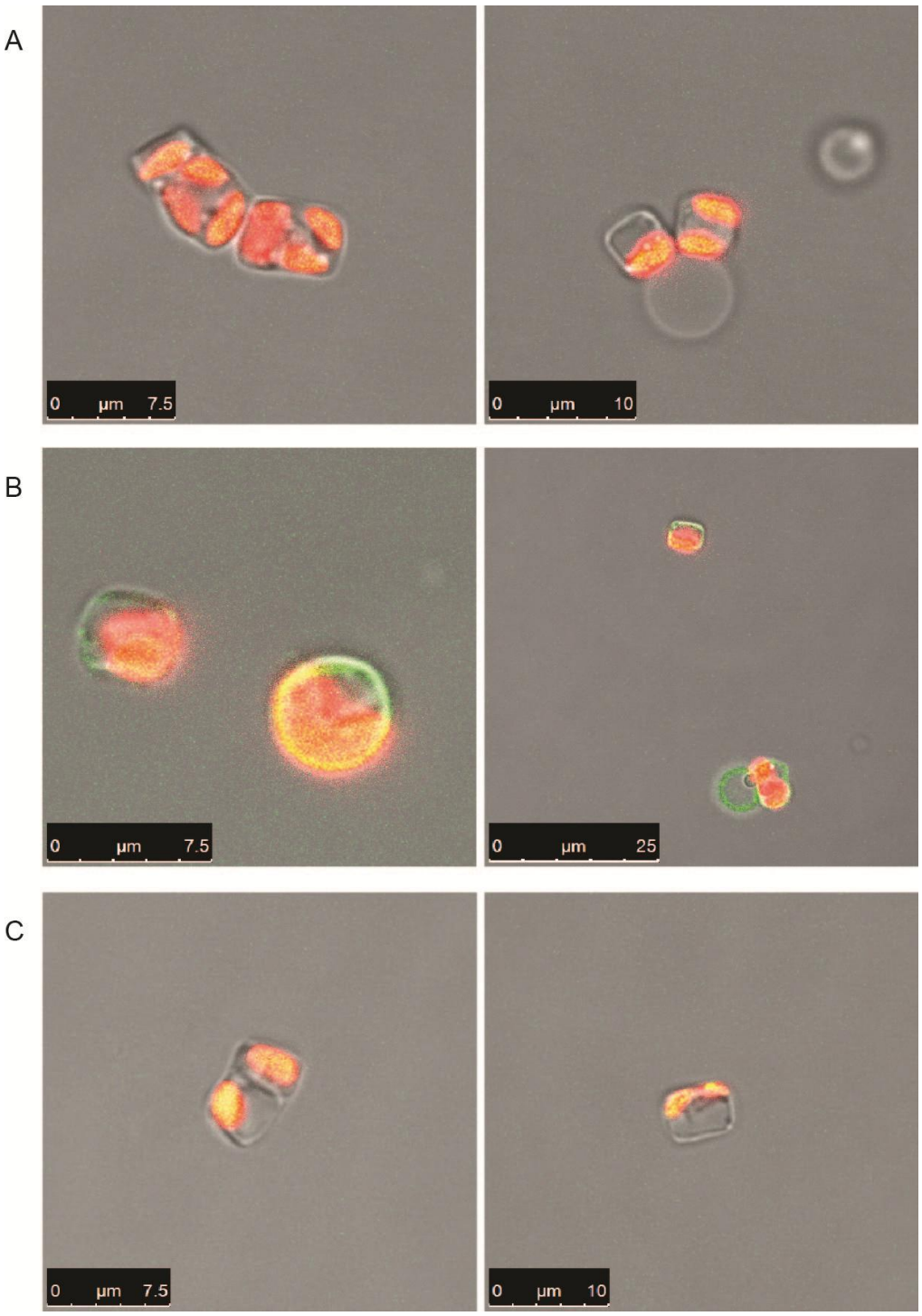


Figure 2.6. Localization of putative HCO_3^- transporters in *T. pseudonana*. A. Tp13887; B. Tp23678; C. Tp270240. Images were collected on Zeiss Axioscope 2 Plus Microscope. Red indicates chlorophyll autofluorescence and yellow or green indicates the YFP fluorescence. Overlap between the green and red shows up as yellow in image B.

CHAPTER 3

SIZE SCALING OF EXTRACELLULAR CARBONIC ANHYDRASE ACTIVITY IN CENTRIC MARINE DIATOMS[2]

[2] Shen, C. and Hopkinson, B.M. 2015. Size scaling of extracellular carbonic anhydrase activity in centric marine diatoms. *Journal of Phycology*. 51: 255-263. dol: 10.1111/jpy.12269. Reprinted here with permission of the publisher.

Abstract

Many microalgae have a surface-associated extracellular carbonic anhydrase (eCA) that converts HCO_3^- to CO_2 for uptake and subsequent photosynthetic fixation. We investigated eCA activity and assessed its importance for photosynthetic CO_2 supply in six centric diatom species spanning nearly the full range of cell sizes for centric diatoms (equivalent spherical radius 3 to 67 μm). Since larger cells are more susceptible to diffusion limitation, we hypothesized that eCA activity would increase with cell size as would its importance for CO_2 supply. eCA activity did increase with cell size, increasing with cell radius by a size-scaling exponent of 2.6 ± 0.3. The rapid increase in eCA activity with cell radius keeps the absolute CO_2 concentration difference between bulk seawater and the cell surface very low (<~0.2 μM) allowing high rates of CO_2 uptake even for large diatoms. Although inhibiting eCA did reduce photosynthesis in the diatoms, there was no overall relationship between the extent of inhibition of photosynthesis and cell size. The only indication that eCA may be more important for larger diatoms was that photosynthesis in the smallest diatoms (<4 μm radius) was only affected by eCA inhibition when CO_2 concentrations were very low, while photosynthesis in some larger diatoms was affected even at typical seawater CO_2 concentrations. eCA is ubiquitous in centric marine diatoms, in contrast to other taxa where its presence is irregularly distributed among different species, and plays an important role in supplying CO_2 for photosynthesis across the size spectrum.

Key index words: carbon uptake; carbonic anhydrase; cell size; diatom; photosynthesis

Abbreviations: ASW, artificial seawater; DBAZ, dextran-bound acetazolamide; DIC, dissolved inorganic carbon; eCA, extracellular carbonic anhydrase; iCA, intracellular carbonic anhydrase; MIMS, membrane inlet mass spectrometry

1. Introduction

Cell size is an important constraint on the acquisition of inorganic carbon and nutrients by phytoplankton. Under nutrient-limited conditions, phytoplankton are capable of taking up the limiting nutrient rapidly enough that its concentration becomes depleted at the cell surface, in which case the rate of nutrient diffusion through the cell surface boundary layer controls uptake (Sunda & Hardison, 2007, Xu et al., 2007). Larger cells are more prone to CO_2 and nutrient limitation because the diffusive flux of these dissolved substances increases only with the cell radius (for spherical or spheroidal cells, which are reasonable approximations for most phytoplankton) while elemental quotas increase with approximately the cube of the cell radius (Pasciak & Gavis, 1974; Morel, 1987; Riebesell et al., 1993). For most nutrients, there is nothing that the cell can do to increase diffusive supply, but CO_2 is a special case because other forms of inorganic carbon, HCO_3^- and CO_3^{2-}, are present in seawater at concentrations that greatly exceed the CO_2 concentration. Although CO_2 is often the preferred substrate for uptake, microalgae can directly take up HCO_3^- or convert HCO_3^- to CO_2 for subsequent uptake (Rost et al., 2003; Martin & Tortell, 2006).

Conversion of HCO_3^- to CO_2 at the surface of phytoplankton cells is catalyzed by a surface-associated extracellular carbonic anhydrase (eCA), a component of the CO_2 concentrating mechanism (CCM). The CO_2 generated can then be taken up for photosynthesis. Alternative roles for eCA in pH homeostasis and recovery of leaked CO_2 have been proposed (Van & Spalding, 1999; Trimborn et al., 2008), but in most cases the primary role for eCA is to supply CO_2 for photosynthesis (Moroney et al., 2011; Hopkinson et al., 2013). eCA activity has been found in a wide range of microalgae including green algae (Fujiwara et al., 1990), diatoms (John-McKay & Colman, 1997), dinoflagellates (Nimer et al. 1999), and haptophytes (Elzenga et

al., 2000). The presence of eCA and its associated activity is often, though not always, regulated by CO_2 availability, being expressed when CO_2 concentrations are low and repressed at high CO_2 concentrations (Nimer et al., 1997). The enzyme has been most well studied in the green algae *Chlamydomonas reinhardtii* where it has been found that two α-CAs in the cell wall or periplasmic space are the major eCAs (Moroney et al., 2011). In the diatom *Thalassiosira pseudonana*, localization of CAs by fluorescent tagging has shown that a ζ-CA and a δ-CA are associated with the cell periphery and are likely the major eCAs (Matsuda et al., 2011; Samukawa et al., 2014). This suggests that eCAs are generally similar in structure and function to other CAs, but are exported from the cell and fixed to the cell surface.

In smaller cells, CO_2 diffusion is rapid enough that even without eCA CO_2 concentrations at the cell surface are not depleted greatly by photosynthetic uptake (Wolf-Gladrow & Riebesell, 1997). Nonetheless some small cells do have eCA in which case the increased CO_2 concentration at the cell surface that eCA produces is thought to provide an energetic benefit for CO_2 uptake (Hopkinson et al., 2013). Large phytoplankton cells on the other hand have the capacity to deplete CO_2 at the cell surface, and may experience CO_2 limitation (Riebesell et al., 1993). For these cells, eCA may be critical to overcome CO_2 limitation and maintain high rates of photosynthesis and growth. In a previous study, we investigated eCA activity in two diatoms: a small diatom, *Thalassiosira pseudonana* (2.5 μm radius) and a larger diatom, *Thalassiosira weissflogii* (6 μm radius) and found that eCA activity was higher in the larger diatom and was more important for photosynthesis, consistent with larger cells being more prone to diffusion limitation (Hopkinson et al., 2013). While this finding was intriguing, it was limited to two diatoms that are only moderately different in size. Here we extend this analysis to cover nearly the full spectrum of centric diatom size from cells of 3 to 67 μm equivalent spherical radius. We

hypothesized that eCA activity would increase with cell size to support greater CO_2 demand by larger cells and that eCA activity would be of greater importance to inorganic carbon supply in larger cells, which are more susceptible to CO_2 limitation. This work takes advantage of a method we recently developed to quantitatively measure eCA activity, which allows direct comparison of diffusive CO_2 fluxes with CO_2 generation rates from eCA (Hopkinson et al., 2013).

2. Materials and Methods

Strains and Culture Conditions

Chaetoceros muelleri Lemmermann (CCMP1316), *Skeletonema marinoi* Sarno & Zingone (CCMP 2092), *Thalassiosira rotula* Meunier (CCMP 1647), *Ditylum brightwellii* (T.West) Grunow (CCMP358), and *Coscinodiscus wailesii* Gran & Angst (CCMP2513), were obtained from the National Center for Marine Algae and Microbiota. *Thalassiosira punctigera* (Castracane) Hasle (CCAP 1085/19) was obtained from Culture Collection of Algae and Protozoa. The algae were grown in artificial seawater medium (ESAW; Harrison et al., 1980) modified as described at the Canadian Center for the Culture of Microorganisms website: (http://www3.botany.ubc.ca/cccm/NEPCC/esaw.html). Manipulation of CO_2 was achieved by control of pH and DIC, rather than bubbling with CO_2-air mixtures, so that the effect of changes in pH, CO_2, and related variables could be independently studied. DIC was omitted from the media initially and later added at the desired concentration (1 or 2.6 mM). EPPS Buffer (5 - 7.5 mM) was added to maintain constant pH, adjusting the initial pH of the culture using HCl or freshly prepared NaOH. All cells were cultured in both a "high CO_2" (initial conditions: pH 7.7, 2.6 mM DIC; CO_2 ~30 μM averaged over growth span) and a "low CO_2" condition (initial conditions: pH 8.4, 2.6 mM DIC; CO_2 ~5 μM averaged over growth span). *D. brightwellii* was

additionally cultured under a "low pH, low CO_2" treatment (initial conditions: pH 7.7, 1 mM DIC; CO_2 ~9 μM averaged over growth span). The initial and final pH of the culture was measured spectrophotometrically at the beginning and end of the experiment (Zhang & Byrne, 1996), and the final DIC concentration was measured by membrane inlet mass spectrometry (MIMS). Over the approximately one week culturing period, pH increased by an average of 0.1 units and an average of 0.4 mM DIC was consumed. All cultures were grown at 18˚C under fluorescent lights (150 μmol photons m^{-2} s^{-1}) on a light/dark cycle (16 hours on, 8 hours off), except *D. brightwellii* and *C. muelleri*, which were grown at 26˚C. The growth of *T. rotula*, *D. brightwellii*, *C. muelleri* and *S. marinoi* was monitored by counting cells daily with a Coulter Counter; *C. wailesii* and *T. punctigera* were monitored by daily chlorophyll fluorescence measurements and the cell numbers were counted under a microscope with Sedgewick-Rafter counting chamber for experimental work. Cells were harvested during exponential growth.

Cell geometry was determined by microscopic observation and measurements with an ocular micrometer for use in determining diffusive fluxes to the cell surface (Table 3.1). The centric diatoms are all approximately cylindrical, but the diffusive equations cannot be exactly solved for a cylinder and so cells were approximated as spheroids (Pasciak & Gavis, 1975). Cells where the height exceeds the diameter were approximated as prolate spheroids with the height dimension as the major axis and radial dimension as the minor axis, whereas cells with diameters greater than heights were approximated with oblate spheroids, taking the radial dimension as the major axis and the height as the minor axis. For spheroids, the shape radius (R_{shape}), which is used in calculating diffusive fluxes to the cell surface, was calculated based on the major and minor axes (Pasciak & Gavis, 1975; Marchetti & Cassar, 2009), and an equivalent spherical radius (R_{sphere}) was calculated from the volume of the cylindrical cell size estimate.

Measurement of eCA activity

The determination of eCA activity first requires measurement of internal carbonic anhydrase (iCA) activity and the mass transfer coefficients for passive CO_2 and HCO_3^- fluxes. These parameters were measured using ^{18}O exchange experiments in the presence of an eCA inhibitor as detailed previously (Hopkinson et al., 2013). In brief, 900 μL ASW/Tris buffer (DIC-free artificial seawater with 20 mM Tris at pH 8.0) was added to the MIMS chamber, followed by 50 μM acetazolamide (AZ), an eCA inhibitor, and 2 mM ^{13}C-^{18}O labeled DIC. AZ, rather than dextran-bound acetazolamide (DBAZ), was used in these experiments because it is somewhat more effective at inhibiting eCA (Hopkinson et al., 2013), and total inhibition of eCA activity is critical for determination of iCA activity and mass transfer coefficients. The temperature in the chamber was maintained at 20˚C using a water jacket. The labeled CO_2 was monitored by MIMS and after ~10 min of monitoring uncatalyzed ^{18}O removal, cells were added from a concentrated suspension. ^{18}O removal by iCA was monitored in the dark for approximately 15 min. The ^{18}O depletion data were then analyzed quantitatively by fitting a model in which iCA activity and the CO_2 and HCO_3^- mass transfer coefficients are unknowns to the data (Tu et al., 1978; Hopkinson et al., 2011). The model treats the interior of the cell as single compartment with iCA activity distributed homogenously throughout the cell. Although diatom CAs are generally not homogeneously distributed (e.g. Tachibana et al., 2011), this simplified representation accurately accounts for the effects of iCA on ^{18}O removal and so can be used to remove the contribution of iCA activity to ^{18}O removal for determination of eCA activity.

eCA activity was then determined from ^{18}O exchange experiments similar to the iCA experiments except that no eCA inhibitor was added. The measurements were made in the dark

at 20 ˚C, using the ASW/Tris buffer described above, and with 2 mM ^{13}C-^{18}O labeled DIC. The ^{18}O depletion data were used to quantify eCA activity by fitting a box model with three compartments: the bulk solution, the cell-surface boundary layer and the intracellular space to the data (Hopkinson et al., 2013). Fitting this model requires knowledge of the iCA activity and CO_2 and HCO_3^- mass transfer coefficients determined as described above, and the uncatalyzed CO_2 hydration and HCO_3^- dehydration rates, which were determined from ^{18}O removal data in the absence of cells. The mass transfer coefficient for diffusive flux between the bulk solution and the boundary layer was calculated as:

$$f_{c-BL} = 4\pi R_{shape} D \quad (1)$$

where R_{shape} is the shape radius of the cell and D is the diffusivity of CO_2 (Pasciak and Gavis, 1975; Marchetti and Cassar, 2009). Only molecular diffusion is considered because turbulent motions are at larger scales than the size of phytoplankton and so will not enhance chemical fluxes (Wolf-Gladrow and Riebesell, 1997). The shape radius is the major axis of the spheroid modified by a shape factor that allows eq. 1 to be used for prolate and oblate spheroids as well as spheres. The first order rate constant for eCA catalyzed CO_2 hydration was determined by optimizing the box model fit to the ^{18}O-CO_2 data. The model fits to the data were good for all species allowing accurate determination of eCA activity (Figure 3.1).

Photosynthetic rates

Photosynthetic rates were determined from the incorporation of ^{14}C-DIC into biomass. Comparing the carbon fixation rates with and without an eCA inhibitor allows an estimate of the effect of eCA activity on photosynthesis. Cells were concentrated during logarithmic growth and re-suspended in an ASW/Tris buffer containing either 2 mM DIC pH 8.2 (typical seawater conditions) or 500 μM DIC pH 8.4 (very low CO_2 conditions). The re-suspended culture was

transferred to acid washed polycarbonate bottles (n=3 for each treatment) and the eCA inhibitor DBAZ was added to a final concentration of 100 μM. DBAZ was used in these experiments rather than AZ to ensure that none of the inhibitor entered the cell, and so that the reductions in photosynthesis are only due to inhibition of eCA. Then 1 μCi (^{14}C-HCO_3) was added and cultures were incubated at 18˚C under fluorescent lights (~200 μmol photons m^2 s^{-1}). The cells were kept in homogeneous suspension by period mixing. After 75 minutes of incubation, 200 μL of culture was removed and mixed with 0.5 mL phenethylamine and scintillation fluid to determine total ^{14}C activity. Cell density was kept very low in the assay requiring a relatively long incubation time. The rest of the sample was filtered onto a 3 μm polycarbonate filter and placed in a scintillation vial. 1 mL 2% HCl was added to remove any residual ^{14}C-DIC and samples were allowed to degas for 1 h. Then 5 mL scintillation fluid was added to all scintillation vials, which were allowed to sit overnight. ^{14}C activity incorporated into biomass was determined by liquid scintillation counting and ^{14}C activity was then converted to photosynthetic rate using the known total C_i concentration, the ^{14}C activity, and incubation time.

Size scaling calculations

In analyzing the size scaling of eCA from a more general perspective it was necessary to estimate net photosynthetic rates (NP) as a function of size and the fraction of nitrogen (N) devoted to eCA. To calculate NP, first the diatom cellular carbon content (Q_c) vs. size relationship of Menden-Deuer and Lessard (2000) was converted from a function of volume to a function of radius assuming spherical geometry, giving $Q_c = 0.077 * R^{2.43}$, where Q_c is in pmol $cell^{-1}$ and R is in μm. The growth rate (μ) vs. size relationship from Wu et al. (2014), which was derived specifically for centric diatoms, was converted to a function of radius, again assuming spherical geometry, giving: $\mu = 1.59 * R^{-0.19}$, where μ is in day^{-1} and R is in μm. Using the fact

that $\mu = NP/Q_c$, a size scaling relationship for NP was derived as: $NP = 0.122 * R^{2.24}$, where NP is in pmol cell^{-1} day^{-1} and R is in μm.

To estimate the fractional N content in eCA as a function of cell size first cellular N content (Q_N) was calculated by dividing Q_c by the Redfield ratio for C:N (6.625). In general the specific eCA proteins have not been identified, but in *T. pseudonana* a well-characterized ζ-CA has been localized to the cell periphery and so this protein was used to estimate eCA N content (Xu et al., 2008; Tachibana et al., 2011). The Zn form of this protein is highly active, at the upper end of the range observed for CAs, so a somewhat lower k_{cat}/K_m of 1 x10^8 $\cdot M^{-1} \cdot s^{-1}$ was used to estimate the moles of eCA required to obtain measured eCA activity (k_{sf}) values (moles eCA per cell = $k_{sf} / (k_{cat} / K_m)$). Moles of enzyme were then converted to moles of N based on the amino acid composition of the *T. pseudonana* ζ-CA and divided by Q_N to estimate the fractional N content devoted to eCA.

3. Results

Extracellular Carbonic Anhydrase (eCA) activity

eCA activity, expressed as the first order rate constant for eCA catalyzed CO_2 hydration (k_{sf}), was measured on the diatom strains cultured under low (~5 μM) and high (~30 μM) CO_2 conditions, which were achieved by manipulating the culture pH (Table 3.2; Figure 3.2). The units of eCA activity ($cm^3 \cdot s^{-1}$) are somewhat unusual being derived from the model framework used to analyze the data; more intuitively the activity can be multiplied by the CO_2 concentration to give the total rate of HCO_3^- generation (in mol $\cdot s^{-1}$) at the cell surface, which at equilibrium is equal to the CO_2 generation rate. eCA activity was detected in all strains, though it was near the limit of detectability in the smallest diatom, *C. muelleri*. In three species (*S. marioni*, *D. brigthwellii*, and *T. punctigera*) eCA activity was up-regulated at low CO_2, with by far the most

dramatic up-regulation observed in *D. brightwellii*. In the other three species (*C. muelleri*, *T. rotula, C. wailesii*) eCA activity was not significantly affected by CO_2 culture condition. To determine if *D. brightwellii*'s increased eCA activity under the "low CO_2" culture condition was actually due to lower CO_2 or the increased pH, *D. brightwellii* was cultured at the same pH of the "high CO_2" condition (7.7), but with reduced DIC (1 mM) to lower the CO_2 concentration (~9 μM). Under this low pH, low CO_2 condition *D. brightwellii*'s eCA activity was high, similar to that under the "low CO_2" culture condition (Figure 3.3), showing that eCA activity was up-regulated in response to low CO_2, not high pH. Among the six strains, eCA activity at low CO_2 closely follows a power law as a function of R_{shape} (non-linear regression: $F_{2,4} = 49.3$, $r^2 = 0.93$, $p = 0.0015$), with a size-scaling exponent of 2.6 ± 0.3 (Figure 3.2).

Effect of eCA inhibition on Photosynthesis

Carbon fixation rates, measured with a ^{14}C method, were quantified with and without the eCA inhibitor dextran-bound acetazolamide (DBAZ) to assess the importance of eCA for carbon acquisition. These measurements were made under two conditions: 1) typical seawater conditions (pH 8.2, 2 mM DIC where CO_2 is 8.5 μM), and 2) a very low CO_2 condition (pH 8.4, 0.5 mM DIC where CO_2 is 1.2 μM). These conditions were based on previous work where it has been found that the effects of eCA inhibition on photosynthesis were only detectable at very low CO_2 concentrations in some phytoplankton (Burkhardt et al. 2001, Hopkinson et al. 2013). Cultures for these experiments were grown under low CO_2 conditions to ensure that eCA was fully expressed. Under typical seawater conditions, inhibition of eCA significantly reduced photosynthetic rates in *D. brightwellii* and *T. rotula* (Table 3.3). At very low CO_2 concentrations, the photosynthetic rates of many more species were significantly reduced by eCA inhibition: *S. marioni*, *D. brigthwellii*, *T. rotula*, *T. punctigera*, and *C. walessii*. Only in the smallest diatom,

C. muelleri, was photosynthesis unaffected by eCA inhibition under either condition. The fractional inhibition of photosynthesis by DBAZ was calculated from the rates of photosynthesis with (P_{DBAZ}) and without ($P_{Control}$) the inhibitor as: $I_{DBAZ}= (P_{Control} - P_{DBAZ})/P_{Control}$ (Table 3.3, Figure 3.4). There was no significant relationship between diatom size and I_{DBAZ} under either assay condition as assessed by linear regression (typical seawater, linear regression: $F_{1,2} = 0.74$, $r^2 = 0.20$, $p = 0.54$, negative value excluded from analysis; very low CO_2, linear regression: $F_{1,3} = 0.01$, $r^2 = 0.00$, $p = 0.93$).

4. Discussion

eCA helps supply CO_2 for photosynthesis by accelerating the dehydration of HCO_3^- to CO_2 at the cell surface. Because larger cells are more prone to diffusive limitation, we hypothesized that larger diatoms would require higher eCA activities and that photosynthesis in larger diatoms would be more dependent on eCA activity.

eCA activity, measured by an ^{18}O isotope exchange technique, increased dramatically with cell size supporting our first hypothesis (Figure 3.2). The best fit to the data was a power law with exponent 2.6 ± 0.3 showing that the increase was more than linear with cell radius. In a previous study (Hopkinson et al. 2013), we presented a framework for interpreting the effectiveness of eCA in terms of the relative supply of CO_2 coming from diffusion through the boundary layer compared with that supplied by eCA-catalyzed dehydration of HCO_3^- at the cell surface. We derived an expression (eq. 2) for the fraction of CO_2 supplied by eCA (E), and equivalently the extent to which eCA reduces the bulk-seawater to cell surface CO_2 gradient:

$$E = \frac{k_{sf}}{f_{c-BL} + k_{sf}} \qquad (2)$$

where k_{sf} is the eCA activity and $f_{c\text{-}BL}$ is the boundary-layer mass transfer coefficient for CO_2. This expression assumes that no CO_2 is generated from dehydration of HCO_3^- within the surface

boundary layer, which is a good assumption for small cells but breaks down as cells get larger (Riebesell et al., 1993). The contribution of CO_2 generated from dehydration of HCO_3^- in the boundary layer can be incorporated into an effective boundary-layer mass transfer coefficient (f'_{c-BL}) based on the analysis of Riebesell et al. (1993):

$$f'_{c-BL} = 4\pi R_{shape} D \left(1 + \frac{R_{shape}}{R_k}\right) \quad (3)$$

where R_{shape} is the shape radius of the spheroidal cells, D is the diffusivity of CO_2, and R_k is reaction-diffusion length scale calculated as $\sqrt{D/k_{uf}}$, where k_{uf} is the CO_2 hydration rate constant. For cells less than 15 μm in radius CO_2 production in the boundary layer is <10% of the diffusive flux under our conditions (18 - 26 ℃, pH 7.7 to 8.4), but for *D. brightwellii* and *T. punctigera* (R_{shape} = 17 and 20 μm respectively) the reactive flux can be up to 12% of the diffusive flux and for *C. wailesii* (R_{shape} = 63 μm) it can be as much as 37%.

Using this modified expression for E, it is clear that E is high for all the diatoms, with the exception of *C. muelleri*, and so eCA is the major source of CO_2 for these diatoms and it significantly reduces the bulk to surface CO_2 gradient (Figure 3.2; Table 3.2). A puzzling aspect of this analysis is that the eCA activity of larger diatoms appears excessive in that a significant reduction in eCA activity (10 fold or even greater) would result in only small decreases in E. This apparent excess can also be seen by comparing the ratio of k_{sf} to f'_{c-BL}, a ratio that would be constant for constant E (Table 3.2). k_{sf} would only need to increase approximately linearly with R (exactly linearly if CO_2 production in the boundary layer is ignored) to maintain a constant effectiveness over relevant cell radii, but k_{sf} was found to increase with a size-scaling exponent of 2.6.

A potential explanation for this additional eCA activity is that it keeps absolute CO_2 gradients between bulk seawater and the cell surface very small across the entire centric diatom

size range. Net photosynthetic (NP) rates in diatoms scale by radius with an exponent of ~2.2 (see Methods). The bulk solution to cell surface CO_2 concentration difference (ΔCO_2) can be calculated from the CO_2 uptake rate (NC) as (Hopkinson et al., 2013):

$$\Delta CO_2 = \frac{NC}{f'_{c-BL} + k_{sf}} \quad (4)$$

Assuming that two-thirds of the C_i required for photosynthesis comes from CO_2 uptake, ΔCO_2 in the absence of eCA activity would need to be as high as ~20 μM for 100 μm cells (Figure 3.5). CO_2 concentrations in the ocean range from ~ 5 μM to 25 μM, but are typically around ~10 μM in temperate waters where CO_2 in the surface ocean is in equilibrium with the atmosphere. Consequently, diffusive limitation could easily limit CO_2 uptake in larger diatoms. ΔCO_2 is reduced by 80% if eCA is present at an activity required to maintain a constant effectiveness of 0.8, the average E value for the diatoms measured here. This level of eCA activity keeps ΔCO_2 below seawater CO_2 concentrations meaning that it is physically possible to support the predicted CO_2 uptake rates, but a substantial ΔCO_2 still develops for larger diatoms (~4 μM at 100 μm cell radius). If instead eCA activity rises with the observed 2.6 scaling exponent, we find that ΔCO_2 remains extremely small (<0.2 μM) across the full range of diatom cell size. The increase in eCA activity with a 2.6 scaling exponent effectively negates the effects of increased CO_2 uptake on ΔCO_2 since the CO_2 uptake rate scales with an exponent of only 2.2.

These results suggest that diatoms express eCA to keep the absolute CO_2 concentration difference between the bulk solution and cell surface low. This is consistent with previous observations that even very small diatoms, which would have only small CO_2 gradients, still produce significant amounts of eCA (Nimer et al., 1998; Hopkinson et al., 2013). CO_2 uptake in the diatom *Phaeodactylum tricornutum* occurs by passive diffusive influx across the cytoplasmic membrane, with the CO_2 deficit in the cytoplasm being generated by the transport of HCO_3^- from

the cytoplasm to the chloroplast (Hopkinson, 2014). CO_2 uptake by passive diffusion is likely to be a common mechanism in diatoms since their cell membranes are highly permeable to CO_2 making active transport by a membrane-embedded transporter a futile process (Hopkinson et al., 2011). Keeping the CO_2 concentration high at the cell surface offers a slight energetic benefit for this uptake mechanism since generating the intracellular CO_2 gradient is an energy consuming process, and so may offer an explanation for the rapid increase in eCA activity with cell size. A rough estimate of the fraction of cellular nitrogen needed for eCA (see Methods) suggests that the cost to produce eCA is relatively small (0.1 - 0.3 % of total cell nitrogen). Consequently, the energetic benefit achieved through reducing the CO_2 gradient in the cell boundary layer may outweigh the energetic and materials costs of producing eCA, particularly for large diatoms.

Decreased CO_2 availability (30 μM to 5 μM) increased eCA activity in three of the diatoms, most dramatically in *D. brigthwellii* where eCA increased ~20x at low CO_2 (Figure 3.2; Table 3.2). However, in two of these species (*S. marioni* and *T. punctigera*) the up-regulation was fairly modest (<2x) and in three species there was no effect of CO_2 culture conditions on eCA activity. This is somewhat at odds with previous work showing that most, though not all, diatoms up-regulate eCA activity dramatically at low CO_2 (Nimer et al., 1997; Trimborn et al., 2008). However, it has also been shown that a moderate sized diatom, *T. weissflogii* (R = 6 μm), down-regulated eCA activity at much higher CO_2 concentrations (>18 μM) than a small diatom, *T. pseudonana* (R=2.5 μm), presumably since larger diatoms require higher CO_2 concentrations to alleviate diffusive-limitation problems (Hopkinson et al., 2013). In this study we examined a number of diatoms larger than *T. weissflogii* for which CO_2 concentrations may need to be substantially higher than our high CO_2 treatment to induce major down-regulation of eCA, potentially accounting for the discrepancy with previous work. Future experiments in which the

culture CO_2 is increased even further may be conducted to resolve this, but care must be taken to avoid lowering the pH to a level that inhibits growth. In the one diatom that did down-regulate eCA activity in our high CO_2 treatment, which was achieved by lowering pH, additional experiments independently manipulating pH and CO_2 availability showed that CO_2 was the key variable regulating eCA activity not pH (Figure 3.3), consistent with previous work on *T. pseudonana* (Hopkinson et al., 2013). This result adds support to the idea that eCA's primary role in diatoms is for CO_2 acquisition.

Our second hypothesis that larger cells would be more dependent on eCA for photosynthesis was generally not supported (Figure 3.4; Table 3.3). Without eCA, larger cells generate greater CO_2 concentration gradients in the surface boundary layer to support the same fraction of photosynthesis by CO_2 uptake (Figure 3.5). Assuming that 1) all diatoms take up CO_2 by the passive process described above and are able to drawdown intracellular CO_2 to similarly low concentrations and 2) that all diatoms use similar fractions of CO_2 and HCO_3^- for photosynthesis, the larger cells will have much lower capacity for CO_2 uptake in the absence of eCA in which case photosynthetic rates would be more negatively affected by eCA inhibition. However, the extent to which photosynthesis was reduced by inhibition of eCA was not related to cell size at either very low CO_2 (0.5 mM DIC; pH 8.4) or typical oceanic conditions (2 mM DIC; pH 8.2) as assessed by linear regression analysis. The only hint of support for this hypothesis was that under typical oceanic conditions photosynthesis was not reduced by eCA inhibition in the smallest diatoms, *S. marioni* and *C. muelleri*, whereas in two of the larger diatoms (*T. rotula*, *D. brightwelli*) it was. Previous work on some small to moderate sized diatoms (*T. pseudonana*, *T. weissflogii*, *Skeletonema costatum*) also found that eCA inhibition only reduced photosynthesis at low CO_2 concentrations, but not at typical oceanic conditions

(Burkhardt et al., 2001; Rost et al., 2003; Hopkinson et al., 2013). It is thus notable that some larger diatoms do require eCA activity to achieve maximal photosynthetic rates under realistic oceanic C_i conditions.

The failure to support this second hypothesis suggests that either the fraction of photosynthesis supported by CO_2 uptake varies significantly among the taxa or that the CO_2 uptake processes differ. The relative contributions of CO_2 and HCO_3^- uptake have indeed been found to vary widely among diatoms, ranging from almost complete CO_2 uptake to complete HCO_3^- uptake (Burkhardt et al., 2001; Trimborn et al., 2008; Martin and Tortell, 2006), and this variability is certainly high enough that it could obscure relationships between eCA dependence and cell size. Alternatively the CO_2 uptake systems may have varying capabilities between the strains. It is likely that some form of passive CO_2 uptake is used by the diatoms, but the nature of the process generating the intracellular CO_2 deficit, and hence the extent to which CO_2 can be drawn down internally, may differ among species. Several diatoms are able to drawdown intracellular CO_2 concentrations to remarkably low levels (<0.5 μM; Hopkinson et al., 2013, Hopkinson, 2014), but other diatoms may not be able to do this and consequently would not be able to take up as much CO_2 without eCA.

The high eCA activity in centric diatoms maintains very low CO_2 gradients between bulk seawater and the cell surface, which likely provides an energetic benefit for CO_2 uptake. Inhibition of eCA results in reduced photosynthesis under typical oceanic conditions in some larger centric diatoms, and under low CO_2 conditions in most of the diatoms. eCA may be especially important during phytoplankton blooms, which diatoms often dominate, when CO_2 concentrations can be reduced significantly below typical concentrations (Takahashi et al., 1993). The ubiquity of eCA in centric diatoms contrasts with other taxa such as haptophytes

(Nimer et al., 1994; Nimer et al., 1997; Rost et al., 2003) and dinoflagellates (Nimer et al., 1999; Rost et al., 2006) where eCA has a more irregular distribution among different species. This distribution suggests that eCA is especially important for CO_2 uptake in centric diatoms and may point to unique characteristics of CCM physiology in this group. Exactly what these characteristics are will have to await more detailed characterization of CCM physiology in these taxa since at present the CCM is only very well characterized in model cyanobacteria (Badger and Price, 2003) and green algae (Moroney and Ynalvez, 2007). Although the CCM is beginning to be fairly well characterized in the model diatoms *Phaeodactylum tricornutum*, and *T. pseudonana* (Reinfelder, 2011; Samukawa et al., 2014).

ACKNOWLEDGEMENTS

This work was supported by NSF grants EF 1041023 and MCB 1129326 to B.M.H.

References

Badger, M. R. & Price, G. D. 2003. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.* 54:609-622.

Burkhardt, S., Amoroso, G., Riebesell, U. & Sültemeyer, D. 2001. CO_2 and HCO_3^- uptake in marine diatoms acclimated to different CO2 concentrations. *Limnol. Oceanogr.* 46:1378-91.

Elzenga, J. T. M., Prins, H. B. A. & Stefels, J. 2000. The role of extracellular carbonic anhydrase activity in inorganic carbon utilization of *Phaeocystis globosa* (Prymnesiophyceae): A comparison with other marine algae using the isotopic disequilibrium technique. *Limnol. Oceanogr.* 45:372-80.

Fujiwara, S., Fukuzawa, H., Tachiki, A. & Miyachi, S. 1990. Structure and differential expression of two genes encoding carbonic anhydrase in *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. U. S. A.* 87:9779-83.

Harrison, P. J., Waters, R. E. & Taylor, F. J. R. 1980. A broad spectrum artificial seawater medium for coastal and open ocean phytoplankton. *J. Phycol.* 16:28-35.

Hopkinson, B. M. 2014. A chloroplast pump model for the CO_2 concentrating mechanism in the diatom *Phaeodactylum tricornutum*. *Photosynth. Res.* 121:223-33.

Hopkinson, B. M., Dupont, C. L., Allen, A. E. & Morel, F. M. M. 2011. Efficiency of the CO_2-concentrating mechanism of diatoms. *Proc. Natl. Acad. Sci. U. S. A.* 108:3830-37.

Hopkinson, B. M., Meile, C. & Shen, C. 2013. Quantification of extracellular carbonic anhydrase activity in two marine diatoms and investigation of its role. *Plant Physiol.* 162:1142-52.

John-McKay, M. E. & Colman, B. 1997. Variation in the occurrence of external carbonic anhydrase among strains of the marine diatom *Phaeodactylum tricornutum* (Bacillariophyceae). *J. Phycol.* 33:988-90.

Marchetti, A. & Cassar, N. 2009. Diatom elemental and morphological changes in response to iron limitation: a brief review with potential paleoceanographic applications. *Geobiology* 7:419-31.

Martin, C. L. & Tortell, P. D. 2006. Bicarbonate transport and extracellular carbonic anhydrase activity in Bering Sea phytoplankton assemblages: Results from isotope disequilibrium experiments. *Limnol. Oceanogr.* 51:2111-21.

Matsuda, Y., Nakajima, K. & Tachibana, M. 2011. Recent progresses on the genetic basis of the regulation of CO_2 acquisition systems in response to CO_2 concentration. *Photosynth. Res.* 109:191-203.

Menden-Deuer, S. & Lessard, E. J. 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* 45:569-79.

Morel, F. M. M. 1987. Kinetics of nutrient uptake and growth in phytoplankton. *J. Phycol.* 23:137-50.

Moroney, J. V., Ma, Y. B., Frey, W. D., Fusilier, K. A., Pham, T. T., Simms, T. A., DiMario, R. J., Yang, J. & Mukherjee, B. 2011. The carbonic anhydrase isoforms of *Chlamydomonas reinhardtii*: intracellular location, expression, and physiological roles. *Photosynth. Res.* 109:133-49.

Moroney, J. V. & Ynalvez, R. A. 2007. Proposed Carbon Dioxide Concentrating Mechanism in *Chlamydomonas reinhardtii*. *Eukaryotic Cell* 6:1251-59.

Nimer, N. A., Brownlee, C. & Merrett, M.J. 1999. Extracellular carbonic anhydrase facilitates carbon dioxide availability for photosynthesis in the marine dinoflagellate *Prorocentrum micans*. *Plant Physiol.* 120:105-11.

Nimer, N. A., Guan, Q. & Merrett, M. J. 1994. Extra-and intra-cellular carbonic anhydrase in relation to culture age in a high-calcifying strain of Emiliania huxleyi Lohmann. *New Phytol*. 126:601-07.

Nimer, N. A., Iglesias-Rodriguez, M. D. & Merrett, M. J. 1997. Bicarbonate utilization by marine phytoplankton species. *J. Phycol.* 33:625-31.

Nimer, N. A., Warren, M. & Merrett, M. J. 1998. The regulation of photosynthetic rate and activation of extracellular carbonic anhydrase under CO_2-limiting conditions in the marine diatom *Skeletonema costatum*. *Plant Cell Environ.* 21:805-12.

Pasciak, W. J. & Gavis, J. 1974. Transport limitation of nutrient uptake in phytoplankton. *Limnol. Oceanogr.* 19:881-98.

Pasciak, W. J. & Gavis, J. 1975. Transport limited nutrient uptake rates in *Ditylum brightwellii*. *Limnol. Oceanogr.* 20:604-17.

Reinfelder, J. R. 2011. Carbon Concentrating Mechanisms in Eukaryotic Marine Phytoplankton. *Ann. Rev. Mar. Sci.* 3:291-315.

Riebesell, U., Wolfgladrow, D. A. & Smetacek, V. 1993. Carbon dioxide limitation of marine phytoplankton growth rates. *Nature.* 361:249-51.

Rost, B., Richter, K., Riebesell, U. & Hansen, P.J. 2006. Inorganic carbon acquisition in red tide dinoflagellates. *Plant Cell Environ.* 29:810-22.

Rost, B., Riebesell, U., Burkhardt, S. & Sültemeyer, D. 2003. Carbon acquisition of bloom-forming marine phytoplankton. *Limnol. Oceanogr.* 48:55-67.

Samukawa, M., Shen, C., Hopkinson, B. M. & Matsuda, Y. 2014. Localization of putative carbonic anhydrases in the marine diatom, *Thalassiosira pseudonana*. *Photosynth. Res.* 121:235-49.

Sunda, W. G. & Hardison, D. R. 2007. Ammonium uptake and growth limitation in marine phytoplankton. *Limnol. Oceanogr.* 52:2496-506.

Tachibana, M., Allen, A. E., Kikutani, S., Endo, Y., Bowler, C. & Matsuda, Y. 2011. Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Photosynth. Res.* 109:205-21.

Takahaski, T, Olafsson J., Goddard, J. G., Chipman, D. W., & Sutherland, S. C. 1993. Seasonal variation of CO_2 and nutrients in the high-latitude surface oceans: a comparative study. *Global Biogeochem.Cy.* 7:843-878.

Trimborn, S., Lundholm, N., Thoms, S., Richter, K., Krock, B., Hansen, P. & Rost, B. 2008. Inorganic carbon acquisition in potentially toxic and non-toxic diatoms: the effect of pH-induced changes in seawater carbonate chemistry. *Physiol. Plant.* 133:92-105.

Tu, C., Wynns, G.C., McMurray, R.E. & Silverman, D.N. 1978. CO_2 kinetics in red cell suspensions measured by ^{18}O exchange. *J. Biol. Chem.* 253:8178-84.

Van, K. & Spalding, M. H. 1999. Periplasmic carbonic anhydrase structural gene (Cah1) mutant in *Chlamydomonas reinhardtii*. *Plant Physiol.* 120:757-64.

Wolf-Gladrow, D. & Riebesell, U. 1997. Diffusion and reactions in the vicinity of plankton: a refined model for inorganic carbon transport. *Mar. Chem.* 59:17-34.

Wu, Y., Campbell, D., Irwin, A. J., Suggett, D. & Finkel, Z. V. 2014. Ocean acidification enhances the growth rate of larger diatoms. *Limnol. Oceanogr.* 59:1027-34.

Xu, Y., Feng, L., Jeffrey, P. D., Shi, Y. & Morel, F. M. M. 2008. Structure and metal exchange in the cadmium carbonic anhydrase of marine diatoms. *Nature.* 452:56-61.

Xu, Y., Tang, D., Shaked, Y. & Morel, F. M. M. 2007. Zinc, cadmium, and cobalt interreplacement and relative use efficiencies in the coccolithophore *Emiliania huxleyi*. *Limnol. Oceanogr.* 52:2294-305.

Zhang, H. N. & Byrne, R. H. 1996. Spectrophotometric pH measurements of surface seawater at in-situ conditions: Absorbance and protonation behavior of thymol blue. *Mar. Chem.* 52:17-25.

Table 3.1. Species and cell geometry. Spheroid indicates the type of spheroid (oblate, prolate) used to approximate the cell. R_1 is the major axis of the spheroid and R_2 is the minor axis of the spheroid. R_{shape} is the shape radius for the spheroid as defined in the text, and R_{sphere} is the equivalent spherical radius of the cell using a cylinder to estimate the cell volume.

Species	Spheroid	R_1 (μm)	R_2 (μm)	R_{shape} (μm)	R_{sphere} (μm)
Chaetoceros muelleri	prolate	3.0	2.4	2.6	3.0
Skeletonema marioni	prolate	4.2	3.5	3.7	4.3
Thalassiosira rotula	oblate	8.0	5.4	7.1	8.0
Ditylum brightwellii	prolate	35	9.5	17	17
Thalassiosira punctigera	oblate	23	15	20	23
Coscinodiscus wailesii	oblate	78	34	63	67

Table 3.2. eCA activity ($cm^3\ s^{-1}$) from the cells acclimated to low and high CO_2 conditions. Values represent the mean of biological replicates and errors are the standard deviation between replicates. E and k_{sf}/f'_{c-BL} were calculated using the eCA activity measured under low CO_2 conditions. eCA activities that were significantly different ($p<0.05$) between CO_2 treatments as assessed by a t-test are indicated with an asterisk.

Species	eCA activity ($cm^3\ s^{-1}$)			
	low CO_2	high CO_2	E	k_{sf}/f'_{c-BL}
Chaetoceros muelleri	$1.6 \pm 0.5 \times 10^{-8}$	$1.8 \pm 1.2 \times 10^{-8}$	0.23	0.3
Skeletonema marioni	$3.5 \pm 0.3 \times 10^{-7}$	$2.4 \pm 0.7 \times 10^{-7}$*	0.82	4.6
Thalassiosira rotula	$2.2 \pm 0.5 \times 10^{-6}$	$2.2 \pm 0.3 \times 10^{-6}$	0.94	15
Ditylum brightwellii	$5.0 \pm 0.4 \times 10^{-6}$	$2.4 \pm 0.4 \times 10^{-7}$*	0.93	14
Thalassiosira punctigera	$3.7 \pm 0.9 \times 10^{-5}$	$2.1 \pm 0.8 \times 10^{-5}$*	0.99	90
Coscinodiscus wailesii	$2.7 \pm 1.5 \times 10^{-4}$	$4.3 \pm 1.2 \times 10^{-4}$	0.99	209

Table 3.3. Carbon fixation rates with and without addition of an eCA inhibitor (DBAZ). The effect of eCA inhibition on photosynthesis was assessed under typical seawater conditions (2 mM C_i, pH = 8.2, CO_2 = 8.5 μM) and at very low CO_2 conditions (0.5 mM C_i, pH = 8.4, CO_2 = 1.2 μM) using cultures grown under low CO_2 conditions. The fractional inhibition of photosynthesis by DBAZ (I_{DBAZ}) was calculated as described in the Results. SD indicates the standard deviation between biological replicates.

	carbon fixation (mol cell^{-1} s^{-1})				
	Control	+DBAZ	I_{DBAZ}	SD	t-test
Chaetoceros muelleri					
very low CO_2	9.07 ± 1.5 x10^{-18}	8.16 ± 0.5 x10^{-18}	0.10	0.17	0.18
typical seawater	1.91 ± 0.8 x10^{-17}	1.34 ± 0.2 x10^{-17}	0.30	0.43	0.14
Skeletonema marioni					
very low CO_2	1.33 ± 0.08 x10^{-17}	7.35 ± 1.3 x10^{-18}	0.45	0.12	0.001
typical seawater	3.13 ± 0.7 x10^{-17}	2.2 ± 0.3 x10^{-17}	0.30	0.27	0.06
Thalassiosira rotula					
very low CO_2	1.97 ± 0.3 x10^{-17}	1.11 ± 0.2 x10^{-17}	0.44	0.20	0.008
typical seawater	5.14 ± 0.5 x10^{-17}	3.32 ± 0.5 x10^{-17}	0.35	0.14	0.005
Ditylum brightwellii					
very low CO_2	3.63 ± 0.5 x10^{-16}	9.71 ± 2.3 x10^{-17}	0.73	0.19	<0.001
typical seawater	9.87 ± 1.2 x10^{-16}	6.5 ± 1.2 x10^{-16}	0.34	0.17	0.01
Thalassiosira punctigera					
very low CO_2	2.57 ± 0.4 x10^{-16}	1.48 ± 0.6 x10^{-16}	0.42	0.29	0.03
typical seawater	1.05 ± 0.2 x10^{-15}	8.51 ± 0.6 x10^{-16}	0.19	0.20	0.08
Coscinodiscus wailesii					
very low CO_2	3.96 ± 0.4 x10^{-15}	2.78 ± 0.3 x10^{-15}	0.30	0.13	0.009
typical seawater	9.05 ± 1.7 x10^{-15}	1.21 ± 0.2 x10^{-14}	-0.34	-0.33	0.07

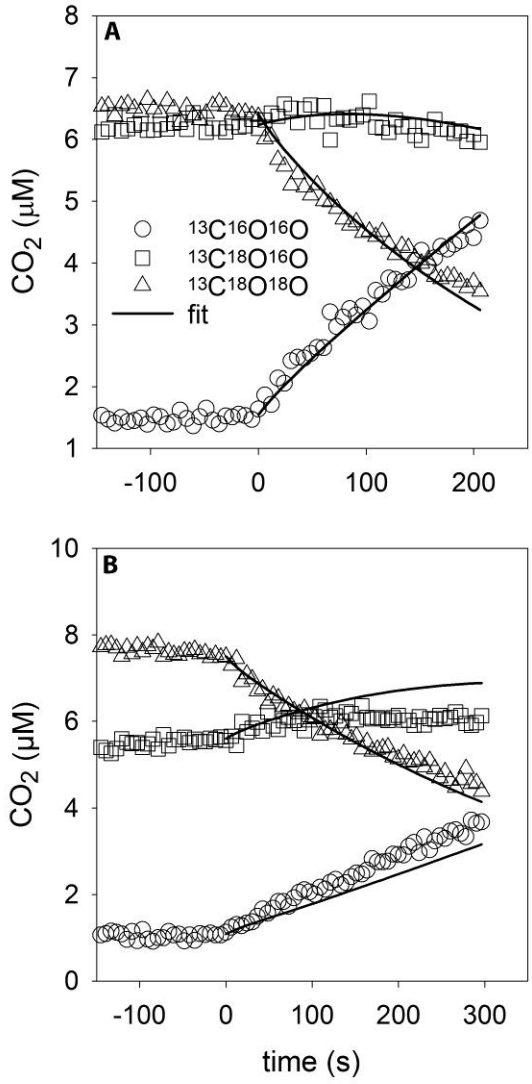


Figure 3.1. Representative results from eCA activity assays showing acceleration of ^{18}O removal from CO_2 upon addition of cells (t=0) and the quality of model fits for (A) *C. wailesii* and (B) *D. brightwellii*. The measurements were made in Tris-buffered ASW (pH 8.0) at 20 ℃ with 2 mM ^{13}C-^{18}O labeled DIC.

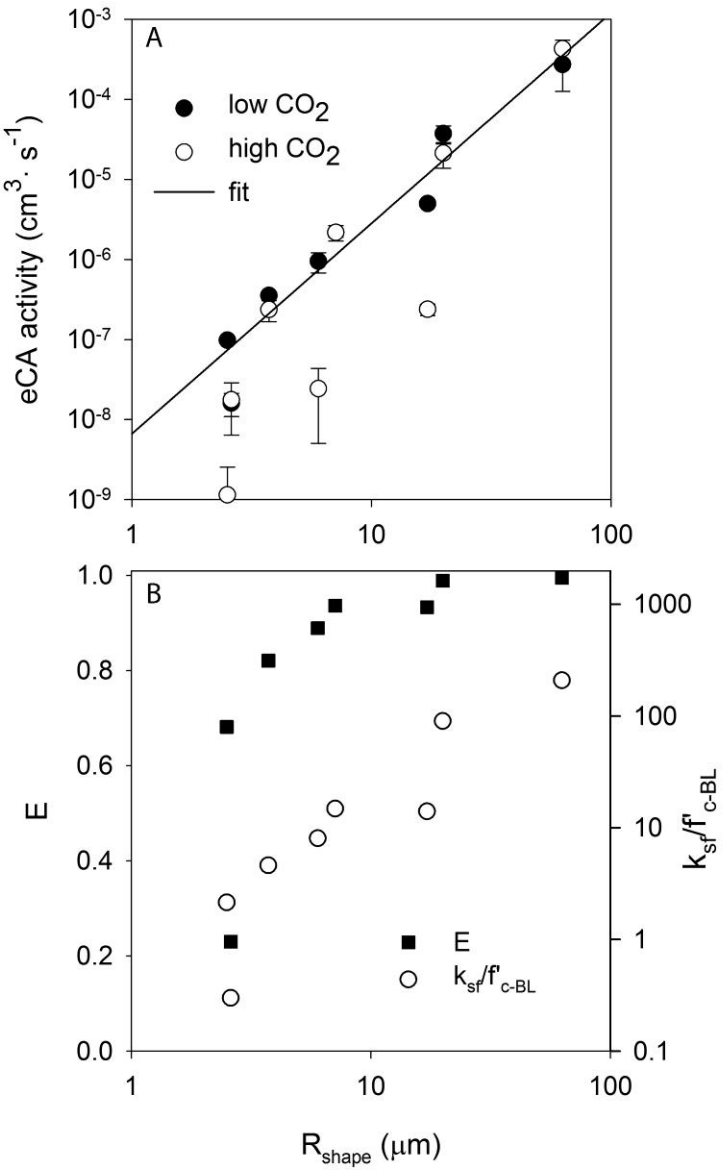


Figure 3.2. (A) eCA activity as a function of shape radius at high and low CO_2. Error bars indicate standard deviations among biological replicates. The line is the power-law best fit to the low CO_2 data (k_{sf}=6.64 x10^{-9} x $(R_{shape})^{2.63}$; R^2 = 0.93). (B) The effectiveness (E) of eCA activity at reducing bulk to cell surface CO_2 concentration gradients and the ratio of eCA activity (k_{sf}) to the boundary-layer mass transfer coefficient for CO_2 (f'_{c-BL}) calculated from low CO_2 eCA activities as a function of the shape radius. Listed by increasing size, the strains are *T. pseudonana*, *C. muelleri*, *S. marioni*, *T. weissflogii*, *T. rotula*, *D. brightwellii*, *T. punctigera* and *C. wailesii*. The data for *T. pseudonana* and *T. weissflogii* was previously published in Hopkinson et al. (2013).

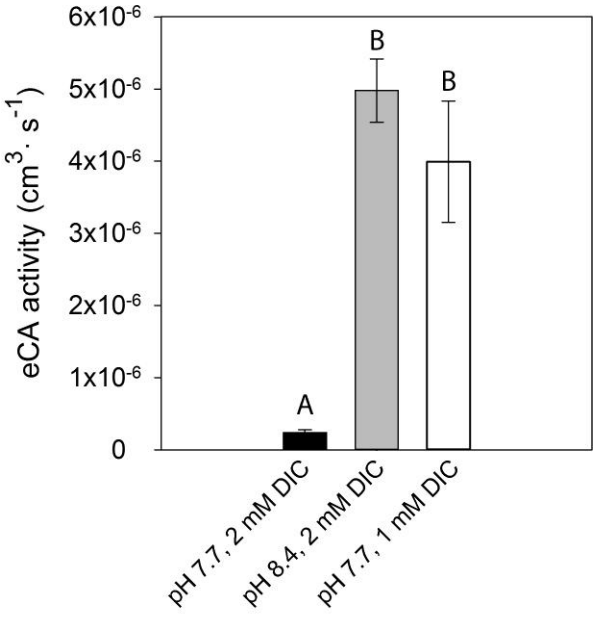


Figure 3.3. eCA activity of *Ditylum brightwellii* cultured under conditions in which DIC and pH were independently varied. Treatments with means that are not significantly different from each other (t-test, $p < 0.05$) are marked with the same letter.

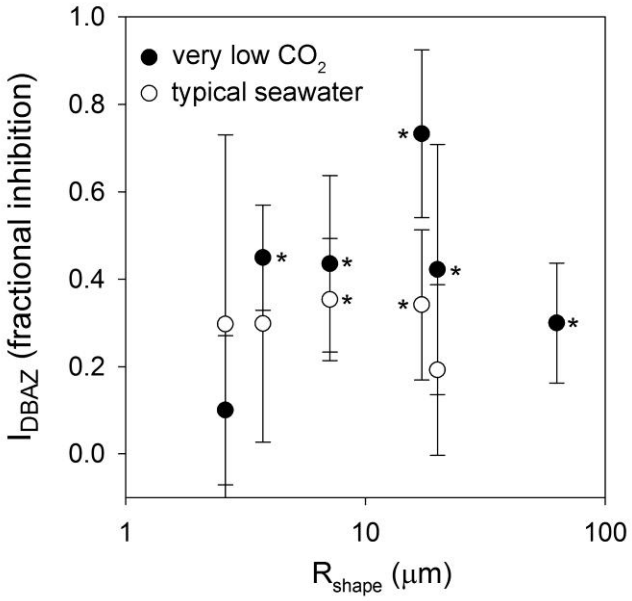


Figure 3.4. Extent of inhibition of carbon fixation by DBAZ, an eCA inhibitor, at two CO_2 concentrations as a function of cell size. Significant inhibition ($p < 0.05$) is indicated with an asterisk. Listed by increasing size, the strains are *C. muelleri*, *S. marioni*, *T. rotula*, *D. brightwellii*, *T. punctigera* and *C. wailesii*. For *C. wailesii*, the typical seawater value was negative and is not shown on the plot.

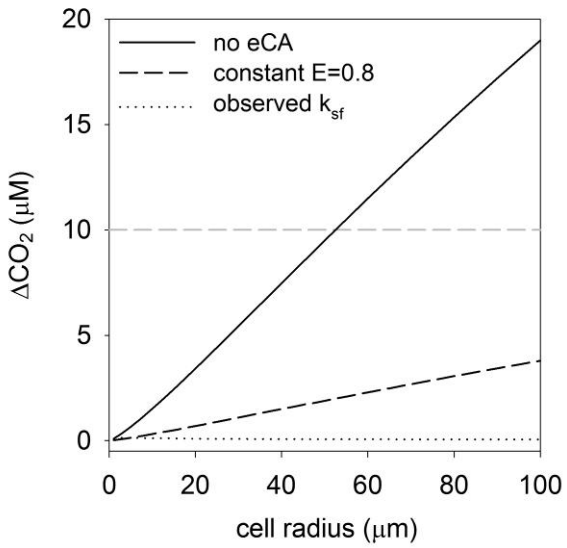


Figure 3.5. Predicted bulk seawater to cell surface CO_2 concentration difference (ΔCO_2) as function of cell radius under different eCA activity scenarios.

CHAPTER 4

DIVERSITY OF CARBON DIOXIDE CONCENTRATING MECHANISMS IN MARINE DIATOMS[3]

[3] Shen, C. and Hopkinson, B. M. To be submitted to *Genome Biology*.

Abstract

Marine diatoms are one of the most ecologically significant primary producers in the ocean. Most diatoms possess a CO_2 concentrating mechanism (CCM) to overcome the scarcity of CO_2 in the ocean and the limitations of the carbon-fixing enzyme RubisCO (Ribulose-1,5-bisphosphate carboxylase/oxygenase). In two model diatoms, *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, several putative CCM components including carbonic anhydrases (CAs) and HCO_3^- transporters have been identified and localized. The different subtypes, locations, and activities of those CCM genes suggest that the CCM differs substantially between these two species. To assess the extent of CCM diversity in marine diatoms more generally, we analyzed genome and transcriptome data from 34 diatom strains to identify putative CCM genes, examine the overall CCM architecture, and study CCM development in the context of the evolutionary history of these diatoms. CCM genes (CAs and HCO_3^- transporters) identified in the diatoms were placed into groups of likely orthologs by OrthoMCL and phylogenetic methods. These analyses indicated that diatoms seem to share similar HCO_3^- transporters, but possess a variety of carbonic anhydrases that have either undergone extensive diversification within the diatom lineage or have been acquired through horizontal gene transfer. Hierarchical clustering of the diatom species based on their CCM gene content suggests that CCM development is largely congruent with evolution of diatom species, despite some notable differences in CCM genes even among closely related species. In particular clustering by CCM gene content separates species with different morphologies that relate to diatom phylogeny. Understanding CCM evolution in the context of past environmental change may shed light on how present day CCMs will response to anthropogenic effects on the environment.

1. Introduction

Diatoms are a group of unicellular photoautotrophic algae that are one of the most ecologically significant primary producers in the ocean (Tréguer et al., 1995; Field et al. 1998; Falkowski et al., 2000). Diatoms can actively take up both CO_2 and HCO_3^- for photosynthesis and this ability to rapidly take up dissolved inorganic carbon (DIC) is critical to their high primary productivity (Trimborn et al., 2008; Matusuda et al., 2011). Due to the high pH and salinity of seawater, dissolved gaseous CO_2 availability is scarce and its diffusion rate is slow, limiting its availability to microalgae (Goyet and Poisson, 1989; Raven et al., 1994; Matsuda et al., 2001). Moreover, the concentration of CO_2 (~10 μM) in modern seawater is not sufficient to saturate rates of carbon fixation by Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), the principal enzyme that catalyzes carbon fixation in the Calvin-Benson cycle (Badger et al., 1998). To overcome the difficulties in acquiring CO_2, diatoms as well as many other marine phytoplankton developed systems called CO_2 concentrating mechanisms (CCMs) to actively take up DIC and increase the CO_2 concentration around RubisCO (Colman and Rotatore, 1995; Johnston and Raven, 1996; Matsuda et al., 2001; Burkhardt et al., 2001; Rost et al., 2003).

Most diatoms use some variety of biophysical CCM, which depends on active pumping of inorganic carbon across cellular membranes (Roberts et al., 2007; Trimborn et al., 2009). Generally, biophysical CCMs consist of CO_2 and HCO_3^- transport mechanisms, intra- and extracellular carbonic anhydrases (CAs), enzymes that catalyze the reversible dehydration of HCO_3^- to CO_2, and a micro-compartment in which RubisCO is concentrated (the pyrenoid in eukaryotes), minimizing the diffusive leakage of CO_2. However, some diatom species such as

Thalassiosira weissflogii are thought to use a biochemical C_4 mechanism, which employs a four carbon organic intermediate for internal translocation (Roberts et al., 2007; Reinfelder, 2000).

Diatoms have a rather complicated evolutionary history including multiple endosymbiotic events, which have involved unknown host cells as well as cyanobacteria, a red alga and possibly a green alga (Moustafa et al., 2009; Kroth, 2002). As a result of a secondary endosymbiotic event, a complicated four-layer chloroplast membrane system was formed, with the outermost membrane linked to the endoplasmic reticulum(ER)-nuclear envelope network (Gibbs, 1981). This compartmentalization makes diatom CCMs more complex since the location of CAs and HCO_3^- transporters can vary among different diatom species (Kroth et al., 2008; Samukawa et al., 2014). There is still no consensus about the ultimate origin of CCMs, but the CCMs in bacteria and eukaryotes show little to no homology, which suggest they evolved independently, and furthermore it is likely that CCMs evolved independently in the major eukaryotic algal lines (Badger et al., 1998 and 2002; Raven et al., 2011 and 2012). Recently, Young et al. (2012) reported that form ID RubisCO in Bacillariophyta (diatoms) and Haptophyta had undergone positive selection during low-CO_2 episodes in geologic history, which possibly relates to the origin of CCMs in these groups.

CCMs in marine diatoms have been most well studied in the model diatoms *Phaeodactylum tricornutum* (a pennate diatom) and *Thalassiosira pseudonana* (a centric diatom) (Tachibana et al., 2011; Samukawa et al., 2014; Nakajima et al., 2013).Ten putative HCO_3^- transporters from solute carrier 4 (SLC4) and solute carrier 26 (SLC26) protein families were found in the *P. tricornutum* genome, and two of these genes have been functionally characterized as HCO_3^- transporters (Nakajima et al. 2013; this thesis). Homologs also exist in *T. pseudonana*, suggesting SLC HCO_3^- transporters might be a common feature in marine diatoms (Nakajima et

al., 2013). No homologs to HCO_3^- transporters from other gene families, such as SbtA from cyanobacteria and LCI1 from the green alga *Chlamydomonas reinhardtii*, have been found in diatom genomes, suggesting that SLC family transporters are the primary mechanism for HCO_3^- transport in diatoms.

Characterization of CAs in these same model diatoms have revealed that while there is some overlap in the families of CAs present in both species (α, γ), β-CAs are only found in *P. tricornutum* and δ-CAs and ζ-CAs are only found in *T. pseudonana* (Tachibana et al., 2011; Samukawa et al., 2014). Even more dramatic differences have been found in the ways in which CAs are distributed throughout the cell in the two diatoms. In *T. pseudonana*, CAs are present in the periplasma, serving to convert HCO_3^- to CO_2 for uptake, whereas no such CAs are found in *P. tricornutum*. Additionally, a CA is found in the cytoplasm of *T. pseudonana*, which likely serves to convert CO_2 diffusing into the cell to HCO_3^-, but *P. tricornutum* lacks cytoplasmic CA. Instead, *P. tricornutum* has multiple CAs within the four-layered chloroplast membrane, which likely serve to convert both CO_2 diffusing into the cell and CO_2 leaking out of the chloroplast into HCO_3^-. *T. pseudonana* has only one CA within the chloroplast membrane sub-compartments. In the chloroplast, *P. tricornutum* lacks CAs in the bulk stroma, allowing an HCO_3^- pool to accumulate, but has two CAs in the pyrenoid, where RubisCO is localized, serving to convert the accumulated HCO_3^- pool to CO_2. In contrast, *T. pseudonana* has a CA distributed throughout the chloroplast stroma, which would be expected to complicate the build up an HCO_3^- pool, suggesting the CCM works quite differently in this diatom (Samukawa et al., 2014).

The dramatic differences in the CCMs of these two model diatoms motivated us to explore CCMs in diatoms more generally, in order to understand the diversity of these systems

and to determine whether the model diatoms really are good representatives of diatoms for studying CCMs. Diatoms have been classified into two major groups, the centrics and the pennates, the names of which refer to their general shape, with centrics having roughly radial symmetry and pennates being elongated, but the division also corresponds to differences in mode of sexual reproduction, plastid number and structure, motility, and other characteristics (Round et al., 1990). Centric diatoms have frequently been divided into two groups of convenience: the radial and polar centrics (Round et al. 1990; Medlin and Kaczmarska, 2004). The radial centrics are mainly circular in outline while the bi- or multipolar diatoms (polar centrics) have non-circular outlines, including triangular and quadrangular forms. Pennates are often divided into two groups as well, the raphid pennates, distinguished by a pair of slits running longitudinally along the frustule (the raphe) that is used in motility, and araphid pennates that lack such structures. The classical view of diatom systematics, mainly based on morphological characteristics, places species into the four taxa described above: radial centrics, polar centrics, raphid pennates, and araphid pennates (Simonsen, 1972 and 1979, Round et al., 1990).

Modern molecular phylogenetic analyses have revealed that these morphological groups have meaningful correspondences with phylogeny, but the details are complicated and still debated (Medlin et al., 1996; Kooistra et al., 2003; Medlin and Kaczmarska, 2004). Medlin and Kaczmarska (2004) identified three monophyletic clades of diatoms: 1) the pennates, 2) polar centrics plus the Thalassiosirales (an order of radial centrics) and 3) the remaining radial centrics (Figure 4.1A). The raphid pennates were also found to be monophyletic, while the araphid pennates were paraphyletic. In this work, the traditional groups described above (radial centrics, polar centrics, raphid pennate, araphid penates) will be used in discussions along with the

monophyletic clades defined by Medlin and Kaczmarska (2004), since they prove useful for discussing the relationship between CCM gene content and diatom phylogeny.

In this work, we examine the CCM gene content of diverse diatoms based on analyses of the 4 available genome sequences and the 30 diatom transcriptomes from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP, Keeling et al., 2014; Figure 4.1B; Table 4.1). We focus on genes that would be involved in both biophysical and biochemical CCMs (CAs and biocarbonate transporters). After grouping CCM genes into putative orthologs, the CCM gene content of the diatoms are compared to each other using hierarchical clustering, and finally the similarity of CCMs among species (based on the hierarchical clustering) is compared to the evolutionary history of diatoms.

2. Materials and Methods

2.1 Data sets

Sets of protein sequences for 34 diatom strains were obtained from the four available genomes (*Thalassiosira pseudonana* CCMP1335; *Phaeodactylum tricornutum*; *Fragilariopsis cylindrus* CCMP1102; *Pseudo-nitzschia multiseries* CLN-47) and from 30 diatom transcriptomes sequenced by the MMETSP (Marine Microbial Eukaryote Transcriptome Sequencing Project, Keeling et al., 2014; Table 4.1). Genomes were obtained from the JGI website (Nordberg et al., 2014) and transcriptomes data were obtained from the iMicrobe website.

2.2 Identification of possible CCM genes in diatom protein sequences

Query sequence sets of HCO_3^- transporters (SLC4) and carbonic anhydrases (α-CAs, γ-CAs and δ-CAs) were collated from the four genomes based on published work (Tachibana et al., 2011; Nakajima et al., 2013; Samukawa et al., 2014) and gene annotation in JGI (Armbrust et al., 2004; Bowler et al., 2008). The query data sets were manually curated and then used to

BLAST against the database of diatom protein sequences with a stringent e-value cut-off of 10^{-5}. HMM searches (Eddy, 1998) were also used to identify possible CAs and HCO_3^- transporters with the e-value cut-off of 10^{-5}. The HMMs were constructed from the same sets of sequences used in the BLAST queries. Sequences that scored below the 10^{-5} cutoff in both the BLAST and HMM analyses were retained. Alignments of the gene families were constructed and used to identify gene fragments and duplicates, which were removed manually.

2.3 Grouping of diatom CCM genes

The putative CA (α-CAs, δ-CAs, and γ-CAs) and HCO_3^- transporter (SLC4) protein sequences identified in the diatoms were grouped into sets using two approaches: 1) OrthoMCL, a program that uses sequence similarity to identify likely orthologs, and 2) a phylogenetic approach using maximum parsimony trees.

OrthoMCL (Chen et al., 2006) was used to cluster components of HCO_3^- transporters (SLC4), α-CAs, δ-CAs, and γ-CAs into groups that represent putative orthologs and 'recent' paralogs. Default parameters were used throughout most of the analysis, but in the final clustering step an inflation parameter of 1.5 was used for HCO_3^- transporters, α-CAs, and γ -CAs components, but the inflation parameter was set to 4 for δ-CAs, which otherwise formed only three or four groups.

A phylogenetic approach was also taken for comparison with the OrthoMCL results. Sequences from each CCM component were aligned by ClustalX 2.1 (Larkin et al., 2007) and the alignments were trimmed using the Gblocks Server (Castresana, 2000; Talavera and Castresana, 2007) and adjusted manually. Maximum parsimony trees were then built using MEGA 6.06 (Tamura et al., 2013) and a bootstrap analysis using 100 resamplings was conducted. The general approach used to define sequence groups was to start from the leaves of the tree and

identify ever larger clades of sequences with greater than 50% bootstrap support. Enlargement of a group was stopped when bootstrap support for more ancestral nodes fell below 50%. Groups with fewer than four sequences were ignored. This approach was used to define most groups. However, in some cases groups were expanded past an internal node whose support fell below 50% when it was deemed to lead to more reasonable groups. Typically this occurred when the poorly supported internal node defined the position of a single sequence relative to a larger clade of sequences. These criteria were developed to ensure the groupings had reasonable phylogenetic support and were at least somewhat comparable to the OrthoMCL groupings (Table 4.2 and Table 4.3).

2.4 Hierarchical Clustering of the diatom species based on CCM gene content

Matrices were complied in which each row represents a diatom species and each column is a CCM gene group defined using either OrthoMCL or phylogenetic methods. The values of each entry in the matrix represent the number of proteins from a CCM gene group found in the diatom species. For consistency with the phylogenetic groups, OrthoMCL groups with less than 4 sequences among all the diatoms were removed. The two matrices, one based on the OrthoMCL groups and the second based on the phylogenetic groups, were used as the input for hierarchical clustering in R (hclust). A similarity matrix was formed based on Euclidean distances and a dendrogram was constructed from the matrix using complete-linkage cluster analysis in R. Comparison of the two hierarchical trees was conducted in R using the package dendextend (Galili, 2015). Comparison of each hierarchical tree with the diatom 18S rDNA phylogenetic tree was done in R using the ape and phytools packages (Paradis et al., 2004; Revell, 2012).

2.5 Diatom phylogenetic tree

A diatom phylogenetic tree was constructed based on the 18S rDNA of the 29 diatoms speices, with *Bolidomonas pacifica* L. Guillou & M.-J. Chretiennot-Dinet used as the outgroup (Genbank ID HQ912557.1). The 18S rDNA sequences were downloaded from the MMETSP website (http://marinemicroeukaryotes.org/resources-files/18s.fa) and Genbank. The 18S rDNA sequence of *C. affinis* was not available. The sequences were aligned using the SILVA Incremental Aligner tool online (SINA, http://www.arb-silva.de/aligner/, Pruesse et al., 2012), which is specifically designed to align rRNA gene sequences. The best nucleotide substation model was found and a maximum likelihood tree was built using Mega 6.06.

3. Results

Four CCM gene datasets (SLC4 HCO_3^- transporters, α-CAs, γ-CAs and δ-CAs) were compiled from the proteins sequences of 34 diatom strains. SLC4 HCO_3^- transporters are presumably the principal HCO_3^- transporters in diatoms (Nakajima et al., 2013). β-CAs and ζ-CAs were not included because these genes were found in very few diatoms. OrthoMCL analysis clustered protein sequences from each family into 5-18 groups (Table 4.2). The HCO_3^- transporters were clustered in five groups and most diatom species had representatives from four of the five groups, indicating that diatoms share similar HCO_3^- transporters (Table 4.2), and conversely that HCO_3^- transporters in a given species come from distinct groups. For example, the nine HCO_3^- transporters from *P. tricornutum* fall into 4 different OrthoMCL groups, and the three HCO_3^- transporters from *T. pseudonana* fell into three different OrthoMCL groups.

In contrast, CAs were generally split into a larger number of groups (8-18) and often the CAs within these OrthoMCL groups were derived from a single diatom genus or even a single diatom species. This trend was most notable in the α-CAs, which were split into 18 groups, with only one group having sequences from more than half of the diatom strains. Among the eight γ-

CAs groups, three groups contained sequences from 32 strains, and most strains possess one γ-CA in each of the three groups. One γ-CA group contained sequences from 20 strains. The remaining groups contained sequences from less than 10 strains. δ-CAs were split into 10 groups, with one group containing δ-CAs from 24 out of 30 strains that possess δ-CAs and the remaining groups only contained sequences from less than 10 strains.

We also built phylogenetic trees with the sequences of each CCM component and used maximal clades with greater than 50% support at most nodes as a criterion for defining a group. The grouping obtained from the HCO_3^- transporters phylogeny showed similar structure to those obtained from OrthoMCL: two clusters covered sequences from most strains (Figure 4.2), indicating diatoms share similar HCO_3^- transporters. On the other hand, phylogenetic trees of CAs tended to form groups within genera or species. For example in the δ-CAs (Figure 4.3), there is a large group containing proteins from *Thalassiosira* and *Skeletonema* species, and smaller groups with sequences from three *Ditylum brightwellii* strains, and from *Pseudo-nitzschia* and *Fragilariopsis* strains. α-CAs and γ-CAs trees also have similar groupings as shown in Figure 4.4 and 4.5. In groups defined by the phylogenetic method, CA sequences from araphid pennates and the radial centrics *Corethron pennatum* L29A3 and *Proboscia alata* PI_D3 often did not fall into the defined groups as a result of low bootstrap support, which may suggest their CAs are quite different from those of other diatoms.

Two independent hierarchical clusterings of diatoms based on their CCM gene content were generated, the first using protein groupings identified using OrthoMCL and the second using protein groupings determined from protein phylogenies. Comparison of the two hierarchical clusterings (Figure 4.6) shows that they are in reasonable agreement, despite some notable exceptions such as the placement of the raphid pennates *Amphiprora sp*. and *Pseudo-*

nitzschia fradulenta WWA7. In both clusterings, the diatom strains are separated into groups that generally correspond to diatom morphology. Raphid pennate diatoms (except *Pseudo-nizschia fradulenta* WWA7 and *Amphiprora sp*. in OrthoMCL clustering) formed a single cluster. Meanwhile there are several clusters formed by centric diatoms, one main cluster containing the order Thalassiosirales: (except *Thalassiosira weissflogii*, *Thalassiosira miniscula* CCMP1093 and *Skeletonema dohrnii* SkelB), and some small clusters composed of polar centrics: one containing all *Chaetoceros* species, and one containing the *Ditylum brightwellii* strains.

The hierarchical clusterings based on CCM gene content were then compared with a diatom species phylogeny (Figure 4.1B) built based on 18S rDNA sequences. Comparing the species phylogeny with the hierarchical clusterings based on the CCM gene content revealed substantial congruence between the two regardless of the method used to define the CCM gene groups (Figure 4.7). Most notably the raphid pennates, which are a monophyletic clade, consistently form a coherent cluster in both hierarchical clusterings (with the exception of *P. fraudulenta* and *Amphiprora sp*. in the OrthoMCL-based clustering). The polar centrics and Thalassiosirales, another monophyletic clade, are mostly contained within a single cluster in the hierarchical clusterings, but other radial centrics and araphid pennates are mixed into this cluster in both the OrthoMCL- and phylogentic-based clusterings. Furthermore, in the phylogenetic-based clustering the *Ditylum brightwelli* strains form a distinct, distant cluster, as do a pair of radial centrics (*S. dohrnii* and *T. miniscula*). The araphid pennates and radial centrics *C. pennatum* and *P. alata* did not show any clear correspondences between their positions in the species phylogeny and hierarchical clusterings.

4. Discussion

CCMs have been studied extensively in cyanobacteria and the green alga *Chlamydomonas reinhardtii*, and have been reasonably well-studied in the model diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* (Price et al, 2008; Jungnick et al., 2014; Tachibana et al., 2011; Nakajima et al., 2013; Samukawa et al., 2014). While the CCMs of different cyanobacterial species are generally similar, the CCM of *Chlamydomonas* differs greatly from that of the cyanobacteria and there is almost no homology between the two systems, implying that they evolved independently (Badger et al., 1998 and 2002; Raven et al., 2011 and 2012). Furthermore the CCMs of *Chlamydomonas* and the model diatoms are quite different, with few components in common (Jungnick et al., 2014). Even the CCMs of the two model diatoms, though making use of many similar components, are organized differently (Samukawa et al., 2014). Despite the paucity of well-characterized CCMs, the little evidence that is available suggests that CCMs, at least within eukaryotes, are subject to extensive diversification leading us to explore the diversity of CCMs within diatoms, an ecologically important group of eukaryotic algae in which the CCM has been reasonably well-studied.

Newly available diatom transcriptomes together with the four sequenced genomes provide a broad sampling of marine diatom taxonomic and environmental diversity (Armbrust et al., 2004; Bowler et al., 2008; Keeling et al., 2014). Most of the sequenced strains were sampled from the Pacific or Atlantic Ocean, but some strains like *Corethron pennatum* L29A3, *Fragilariopsis kerguelensis* L2_C3, and *Thalassiothrix Antarctica* L6_D1 were sampled from the Southern Ocean. The diatoms are disproportionally from coastal waters, but many came from open-ocean environments, and several were obtained from estuaries (e.g. *Skeletonema marinoi* SkelA, *Skeletonema dohrnii* SkelB and *Thalassionema nitzschioides* L26_B). Our dataset includes the four major diatom morphologies, with 8 polar centrics, 14 radial centrics (12

belonging to the order Thalassiosirales), 9 raphid pennates and 3 araphid pennates (Figure 4.1B, Table 4.1).

The primary components of the diatom CCM identified to date are SLC4 HCO_3^- transporters and several groups of carbonic anhydrases, of which the α-, γ-, and δ-CAs families are widespread among diatoms (Roberts et al., 1997; Lane and Morel, 2000; Tachibana et al., 2011; Nakajima et al., 2013; Samukawa et al., 2014). These CCM components were identified in the transcriptomes and genomes of diatoms using BLAST and HMM analyses. A limitation of the data set is the large number of transcriptomes since not all genes will necessarily be expressed under the particular culturing conditions used to generate the data. We compared the CCM components found in the *P. tricornutum* genome with those present in a *P. tricornutum* transcriptome (C. Dupont, unpublished) and found that 16 of the 18 genes found in the genome were also expressed in the transcriptomes. This provides some confidence that the CCM genes are generally expressed and our analyses are not grossly biased.

After compiling CCM protein sequences, sequences from each of the four CCM components (SLC4 HCO_3^- transporters, α-CAs, γ-CAs, and δ-CAs) were grouped using OrthoMCL and phylogenetic trees into groups of related sequences that ideally play a similar functional role in the CCM. HCO_3^- transporters were grouped by both methods into a few major groups that contain sequences from most diatom strains (Figure 4.2, Table 4.2), suggesting that nearly all diatoms contain a similar set of HCO_3^- transporters that likely function in different roles in the cell. For example, one group may be localized to the plasma membrane bringing HCO_3^- into the cell (Nakajima et al., 2013), while another group may be embedded in the chloroplast membranes transporting HCO_3^- into the chloroplast. In addition, HCO_3^- transporters from *T. pseudonana* and *P. tricornutum* were distributed throughout most of the major groups,

which demonstrates that the CCMs of these model diatoms are in this respect representative of diatoms as a group.

In contrast to HCO_3^- transporters, the α-CAs and δ-CAs only had one sequence group that contained sequences from most diatom strains. The remaining sequence groups for these CAs were comprised of sequences with specific taxonomic affiliations (Figure 4.3 and 4.4). In some cases these were broad taxonomic groups (e.g. pennates Figure 4.3B and 4.4), but in other cases there groups were composed entirely of sequences from a single genus or even species. The most extreme examples of this are found in *D. brightwellii*, where both α-CAs and δ-CAs appear to have undergone extensive radiation (Figure 4.3B and 4.4). OrthoMCL put γ-CA sequences into several clean groups that contained sequences from most strains. Phylogenetic methods instead separated the γ-CAs along lines reflecting species phylogeny, but close inspection of the γ-CA tree indicates this is in part an artifact of the criterion used to define groups (Figure 4.5A). In general, the CAs typically have one or two groups from raphid pennate diatoms (e.g. *Pseudo-nizschia* and *Fragilariopsis* species), some groups composed of genes from *Thalassiosira* and *Skeletonema* species, *Chaetoceros* species, and a separate group composed of sequences from the three *Ditylum brightwellii* strains.

Overall, analysis of CA affiliations indicated they are diverse and differentiated within diatoms, which in turn indicates that CAs have evolved rapidly within the diatom lineage, or perhaps that they have been acquired through horizontal gene transfer at different stages of their evolution. This perspective is also supported by experimental evidence, showing the differences of CA types, locations, and activities in *T. pseudonana* and *P. tricornutum* (Tachibana et al., 2011; Samukawa et al., 2014). Five α-CAs have been localized to chloroplast membrane system in *P. tricornutum* and presumably function to control DIC flux into and out of the chloroplast,

while in *T. pseudonana*, there is only one α-CAs which was localized in the chloroplast stroma. δ-CAs are not found in *P. tricornutum* but there are four in *T. pseudonana* distributed throughout the cell (on the surface, in the mitochondria, and in the chloroplast membrane system) where they fulfill different roles (Tachibana et al., 2011; Samukawa et al., 2014).

In general, CAs from these two model diatoms showed distinct subtypes and locations, which suggest different roles in CCM and result in a variety of CCM architectures. However, the γ-CAs are a notable exception. Two γ-CAs from *P. tricornutum* and three γ-CAs from *T. pseudonana* were all found to localize to the mitochondrion (Tachibana et al., 2011; Samukawa et al., 2014). Mitochondrial localization of γ-CAs was also found in the higher plant *Arabidopsis thaliana* (Parisi et al., 2004), but these putative γ-CAs lack CA activity, suggesting some γ-CAs might have different functions (Klodmann et al., 2010). Phylogenetic analysis of diatom γ-CAs showed diatom γ-CAs were distant from those of *A. thaliana* (data not shown; Tachibana et al. 2011). Within the diatoms, the three confirmed mitochondrion γ-CAs from *T. pseudonana* and one predicted mitochondrial γ-CA fall into four separate groups (Figure 4.5; γ-CAs from *P. tricornutum* did not fall in any defined groups). Similarly in OrthoMCL groups, mitochondrial γ-CAs from both *T. pseudonana* and *P. tricornutum* fall into separate groups, each of which contained sequences from most strains. The consistent localization of multiple γ-CAs to the mitochondrion in both diatoms and higher plants suggests that they play an important, functionally-conserved role in these organisms. What that role is remains unclear, but notably a γ-CA subcomplex is found in the mitochondrial complex I of photoautotrophic eukaryotes, such as green alga and plants, but not in that of the heterotrophic eukaryotes like fungi and mammals, suggesting these γ-CAs are related to photosynthetic carbon metabolism (Klodmann et al., 2010; Hunte et al., 2010).

Hierarchical clustering was used to assess relationships among diatoms based on their CCM gene content, and clusterings based on CCM structure were then compared with species phylogeny to assess the extent to which diatom CCMs have diversified or remained conserved within phylogenetic lines. Two hierarchical clusterings built from matrices based on CCM gene content were generated, one using protein groups defined with OrthoMCL and the second using protein groups defined by phylogenetic analysis (Figure 4.6). The two clusterings are generally in agreement, and both show certain consistent correspondences between CCM structure and diatom morphology: raphid pennates form a coherent group and centric diatoms generally cluster together, but the radial centrics (*Proboscia alata* PI_D3) and araphid pennates (*Asterionellopsis glacialis* CCMP134, *Thalassiothrix Antarctica* L6_D1, *Thalassionema nitzschioides* L26_B) are mixed in, apparently haphazardly, with the centrics cluster.

The relationship between CCM structure and diatom morphology, which roughly reflects phylogeny, motivated us to analyze the relationship between diatom species evolution and CCM structure in more detail. A diatom species phylogeny was built based on 18S rDNA sequences. Though considerable advances have been made with molecular phylogenetic approaches, diatom phylogenetics are still a topic of debate, particularly regarding higher-level relationships in diatoms, among families, orders, and classes (Medlin and Kaczmarska 2004; Sims et al. 2006; Theriot et al., 2009). The major divisions found in our 18S phylogeny mostly supported Medlin and Kaczmarska phylogeny schema (Medlin and Kaczmarska 2004) in which three major clades are found: pennates, polar centrics and Thalassiosirales, and the remaining radial centrics (Figure 4.1).

The 18S rDNA species phylogeny and the two hierarchical clusterings of species based on CCM gene content show substantial agreement (Figure 4.7), indicating that diatom CCM

development has largely proceeded along phylogentic lines rather than being dominated by horizontal gene transfer. Nonetheless, the CCM has diversified substantially within diatom lineages and there is variation between the species phylogeny and clusterings based on CCM structure. The most notable concordance is among the raphid pennates, which form consistent groups in the CCM-based clustering (with the exception of *P. fraudulent* and *Amphipohora sp* in the OrthoMCL-based clustering) and are a monophyletic taxon. Relationships within the raphid pennates are also mostly congruent between the species phylogeny and CCM clusterings. The polar centrics and Thalassiosirales also show substantial congruence between the species phylogeny and CCM clusterings, but there are some interesting disparities. *D. brightwellii* belongs to the order Lithodesmiales, which is a sister group of Thalassiosirales based on SSU rRNA (Medlin and Kaczmarska, 2004), as shown in the 18S rDNA phylogenetic tree (Figure 4.1B). However, the *D. brightwellii* strains form distinct grouping based on CCM structure, driven by diversification of CAs within this species. The remaining radial centrics and araphid pennates are distributed throughout the CCM clusterings with no obvious correspondence to their phylogenetic positions. In summary, clearly diatom CCMs are diverse and there is evidence for both gradual developments of CCMs during the evolution of diatom species and for anomalous CCM development in certain diatoms.

Insight into potential factors driving CCM diversification is found in the evolutionary history of diatoms. Fossil evidence indicates that eukaryotic marine algae originated around 1.2 Gya (Butterfield, 2000 and 2004; Butterfiled et al., 1988) and molecular phylogenetic analysis suggests RubisCO evolved before the origin of oxygenic photosynthesis (Tabita et al., 2007, 2008a, 2008b). The majority of photoautotrophs that have evolved since that time use RubisCO for photosynthetic carbon fixation (Hohmann-Marriott and Blankenship, 2011). Diffusive supply

of CO_2 to RubisCO was presumably the ancestral mechanism of carbon supply in oxygenic photoautotrophs since CO_2 concentrations were high and O_2 was absent early in earth's history. But as CO_2 decreased and O_2 increased, reliance on diffusive CO_2 supply would begin to result in lower rates of fixation by RubisCO and increased rates of photorespiration. To overcome these inefficiencies, the CCM appeared as an important evolutionary response to maintain photosynthetic performance (Raven et al., 2011 and 2012). There were several periods of low CO_2 during which CCMs could have evolved, including the Permo-Carboniferous glaciations 320-270 Ma, and repeated glacial intervals of the Pleistocene 2.1 Ma (Raven et al., 2011). Although there is no direct fossil evidence or molecular clocks to show the origin of CCMs, positive selection on form ID RubisCO in Bacillariophyta (diatoms) and Haptophyta during periods of low-CO_2 could relate to the origin or reengineering of CCMs in these taxa (Young et al., 2012). These repeated oscillations of periods of low and high CO_2 could induce periodic pressure to develop or redesign the CCM, leading to continual diversification of the CCM as observed in our analysis of diatom CCM gene content.

If the cyanobacterial CCM and some algal CCMs evolved before the Pleistocene after which CO_2 concentrations have remained low, the question arises as to why CCMs were retained by algae during periods of high CO_2 and high temperatures. The retention of CCMs in these unfavorable environments in the past may shed light on how present day CCMs will respond to future conditions in which CO_2 and temperature rise. Raven et al. (2005, 2011, and 2012) summarized the effects of several environmental factors on the extent of CCM engagement. Based on observations of extant algae to different environment conditions, Raven et al. (2012) suggested the retention of CCMs in high CO_2 episodes could have been related partially to the interaction of CCMs with other environmental factors like photosynthetically active radiation

(PAR), ultraviolet-B radiation (UV-B) and nutrients (Giordano et al. 2005; Raven et al. 2005, 2008; Beardall et al. 2009a, c; Gao et al. 2009). Under high CO_2 and high temperature conditions, DIC and H^+ concentration will increase in surface seawater, while warming leads to shallower upper mixed layers. As a result, availability of inorganic carbon, PAR and UVB increases while supply of nutrients like N and P decreases. Short-term experiments on extant phytoplankton reveal that the decreased nutrient supply, with the increased mean flux of PAR and UV-B will lead to more involvement of CCMs, creating selective pressure for their retention even under high CO_2 conditions (Raven et al., 2011 and 2012). Consequently, while CO_2 concentrations are likely to be the primary factor driving CCM development, other environmental factors may also exert some selective pressure.

5. Conclusions

Analyzing HCO_3^- transporters and carbonic anhydrases, the key components of diatom CCMs, in the genomes and transcriptomes of 34 diatom strains revealed a great diversity of CCM architecture within diatoms. Much of this diversity corresponds with diatom species phylogeny, but the CCM has diverged substantially in some lineages. While SLC4 HCO_3^- transporters are generally similar among diatom species, there has been extensive development of α- and δ-CAs within certain taxonomic groups. These results show that the model diatoms *P. tricornutum* and *T. pseudonana* do not capture the full diversity of diatom CCMs. Although Medlin and Kaczmarska's hypothesis of diatom phylogeny has been debated, our 18S phylogeny mostly supports the major divisions. Furthermore, the CCM structure was characterized here confirms the special position of Thalassiosirales. However, our dataset contained only 30 transcriptomes and characterization of CCM components is still under study in these diatoms. A

thorough analysis of CCM diversity and evolution would ideally use more genomic data and be based on more detailed knowledge of CCM components.

References

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., Zhou, S., Allen, A. E., Apt, K. E., Bechner, M., Brzezinski, M. A., Chaal, B. K., Chiovitti, A., Davis, A. K., Demarest, M. S., Detter, J. C., Glavina, T., Goodstein, D., Hadi, M. Z., Hellsten, U., Hildebrand, M., Jenkins, B. D., Jurka, J., Kapitonov, V. V., Kröger, N., Lau, W. W., Lane, T. W., Larimer, F. W., Lippmeier, J. C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M. S., Palenik, B., Pazour, G. J., Richardson, P. M., Rynearson, T. A., Saito, M. A., Schwartz, D. C., Thamatrakoln, K., Valentin, K., Vardi, A., Wilkerson, F. P., Rokhsar, D. S. 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*. 306: 79-86.

Alverson, A. J. and Theriot, E. C. 2005. Comments on recent progress toward reconstructing the diatom phylogeny. *J. Nanosci. Nanotechnol*. 5: 57-62.

Badger, M. R., Andrews, T. J., Whitney, S. M., Ludwig, M., Yellowlees, D. C., Leggat, W., Price, G. D. 1998. The diversity and coevolution of rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. *Canadian Journal of Botany*. 76: 1052-1071.

Badger, M. R., Hanson, D., Price, G. D. 2002. Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. *Funct Plant Biol*. 29: 161-173.

Beardall, J., Sobrino, S., Stojkovic, S. 2009a. Interactions between impacts of ultraviolet radiation, elevated CO2, and nutrient limitation in marine primary producers. *Photochem Photobiol Sci*. 8: 1257-1265.

Beardall, J., Stojkovic, S., Larson, S. 2009c. Living in a high CO2 world; impacts of global climate change on marine phytoplankton. *Plant Ecol Divers*. 2: 191-205.

Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R. P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J. A., Brownlee, C., Cadoret, J. P., Chiovitti, A., Choi, C. J., Coesel, S., De Martino, A., Detter, J. C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman, M. J., Jenkins, B. D., Jiroutova, K., Jorgensen, R. E., Joubert, Y., Kaplan, A., Kröger, N., Kroth, P. G., La Roche, J., Lindquist, E., Lommer, M., MartinJezequel, V., Lopez, P. J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L. K., Montsant, A., Oudot-Le Secq, M. P., Napoli, C., Obornik, M., Parker, M. S., Petit, J. L., Porcel, B. M., Poulsen, N., Robison, M., Rychlewski, L., Rynearson, T. A., Schmutz, J., Shapiro, H., Siaut, M., Stanley, M., Sussman, M. R., Taylor, A. R., Vardi, A., von Dassow, P., Vyverman, W.,

Willis, A., Wyrwicz, L. S., Rokhsar, D. S., Weissenbach, J., Armbrust, E. V., Green, B. R., Van de Peer, Y., Grigoriev, I. V. 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*. 456: 239-244.

Burkhardt, S., Amoroso, G., Riebesell, U., Sültemeyer, D. 2001. CO_2 and HCO_3^- uptake in marine diatoms acclimated to different CO_2 concentrations. *Limnol Oceanogr*. 46: 1378-1391.

Butterfield, N. J. 2000. Bangiomorpha n. gen., n. sp.: implications for the evolution of sex, unicellularity, and the Mesoproterozoic/Neoproterozic radiation of eukaryotes. *Palaeobiology*. 26: 386-404.

Butterfield, N. J. 2004. A vaucheriacean alga from the middle Neoproterozoic of Spitsbergen: implication for the evolution of Proterozoic eukaryotes and the Cambrian explosion. *Palaeobiology*. 30: 231-252.

Butterfield, N. J., Knoll, A. H. and Swett, K. 1988. Exceptional preservation of fossils in an upper Proterozoic shale. *Nature*. 334: 424-427.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*. 17: 540-552.

Chen, F., Mackey, A. J., Stoeckert, C. J. Jr., and Roos, D. S. 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res*. 34: D363-8.

Colman, B. and Rotatore, C. 1995. Photosynthetic inorganic carbon uptake and accumulation in two marine diatoms. *Plant Cell Env*. 18: 919-924.

Eddy, S. R. 1998. Profile hidden Markov models. *Bioinformatics*. 14(9): 755-763.

Falkowski, P., Scholes, R. J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Hoegberg, P., Linder, S., Mackenzie, F. T., Moore III, B., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., Steffen, W. 2000. The global carbon cycle: a test of our knowledge of Earth as a system. *Science*. 290: 291-296.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., Falkowski, P. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. 281: 237-240.

Galili, T. 2015. dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics*. 31(22): 3718-3720.

Gao, K. S., Ruan, Z. X., Villafane, V. E., Gattuso, J. P., Helbling, E. W. 2009. Ocean acidification exacerbates the effect of UV radiation on the calcifying phytoplankton *Emiliana huxleyi*. *Limnol Oceanogr*. 54: 1855-1862.

Gibbs, S. P. 1981. The chloroplast endoplasmic reticulum: structure, function and evolutionary significance. *Int Rev Cytol*. 72: 49-99.

Giordano, M., Beardall, J., Raven, J. A. 2005. CO2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu Rev Plant Biol.* 56: 99-131.

Goyet, C., Poisson, A. 1989. New determination of carbonic acid dissociation constants in seawater as a function of temperature and salinity. *Deep-Sea Res*. 36: 1635-1654.

Hohmann-Marriott, M. F. and Blankenship, R. E. 2011. Evolution of photosynthesis. *Annu. Rev. Plant Biol*. 441: 940-941.

Hunte, C., Zickermann, V., Brandt, U. 2010. Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science*. 329: 448-451.

Johnston, A. M., Raven, J. A. 1996. Inorganic carbon accumulation by the marine diatom Phaeodactylum tricornutum. *Eur J Phycol* 31: 285-290.

Jungnick, N., Ma, Y., Mukherjee, B., Cronan, J. C., Speed, D. J., Laborde, S. M., Longstreth, D. J., Moroney, J. V. 2014. The carbon concentrating mechanism in *Chlamydomonas reinhardtii*: finding the missing pieces. *Photosynth Res*. 121: 159-173.

Kaczmarska, I., Beaton, M. and Benoit, A.C. 2005. Molecular phylogeny of selected members of the order Thalassiosirales (Bacillariophyta) and evolution of the fultoportula. *J. phycol*. 42: 121-138.

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., Armbrust, E. V., Archibald, J. M., Bharti, A. K., Bell, C. J., Beszteri, B., Bidle, K. D., Cameron, C. T., Campbell, L., Caron, D. A., Cattolico, R. A., Collier, J. L., Coyne, K., Davy, S. K., Deschamps, P., Dyhrman, S. T., Edvardsen, B., Gates, R. D., Gobler, C. J., Greenwood, S. J., Guida, S. M., Jacobi, J. L., Jakobsen, K. S., James, E. R., Jenkins, B., John, U., Johnson, M. D., Juhl, A. R., Kamp, A., Katz, L. A., Kiene, R., Kudryavtsev, A., Leander, B. S., Lin, S., Lovejoy, C., Lynn, D., Marchetti, A., McManus, G., Nedelcu, A. M., Menden-Deuer, S., Miceli, C., Mock, T., Montresor, M , Moran, M. A., Murray, S., Nadathur, G., Nagai, S., Ngam, P. B., Palenik, B., Pawlowski, J., Petroni, G., Piganeau, G., Posewitz, M. C., Rengefors, K., Romano, G., Rumpho, M. E., Rynearson, T. A., Schilling, K. B., Schroeder, D., Simpson, A. G. B., Slamovits, C. H., Smith, D. R., Smith, G. J., Smith, S. R., Sosik, H. M., Stief, P., Theriot, E. C., Twary, S. N., Umale, P. E., Vaulot, D., Wawrik, B., Wheeler, G., Wilson, W. H., Xu, Y., Zingone, A. and Worden, A. Z. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLoS Biology*. 12(6): e1001889.

Klodmann, J., Sunderhaus, S., Nimtz, M., Jansch, L., Braun, H. P. 2010. Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. *Plant Cell*. 22: 797-810.

Kooistra, W. H. C. F., De Stefano, M., Mann, D. G., Salma, N. & Medlin, L. K. 2003. Phylogenetic position of Toxarium, a pennate-like lineage within centric diatoms (Bacillariophyceae). *J. Phycol*. 39: 185-97.

Kroth, P. G. 2002. Protein transport into secondary plastids and the evolution of primary and secondary plastids. *Int Rev Cytol*. 221:191–255.

Kroth, P. G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M. S., Stanley, M. S., Kaplan, A., Caron, L., Weber, T., Maheswari, U., Armbrust, E. V., Bowler, C. 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS One*. 3: e1426.

Lane, T. W., Morel, F. M. M. 2000. A biological function for cadmium in marine diatoms. Proc *Natl Acad Sci USA*. 97: 4627-4631.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. and Higgins*, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*. 23(21): 2947-2948.

Mann, D. G. 2010. Diatoms. Version 07 February 2010 (under construction). http://tolweb.org/Diatoms/21810/2010.02.07 in The Tree of Life Web Project, http://tolweb.org/

Matsuda, Y., Hara, T., Colman, B. 2001. Regulation of the induction of bicarbonate uptake by dissolved CO_2 in the marine diatom, *Phaeodactylum tricornutum*. *Plant Cell Env*. 24: 611-620.

Matsuda, Y., Nakajima, K., Tachibana, M. 2011. Recent progresses on the genetic basis of the regulation of CO_2 acquisition systems in response to CO_2 concentration. *Photosynth Res*. 109: 191-203. Medlin, L. K., Williams, D. M. and Sims, P. A. 1993. The evolution of the diatoms (Bacillariophyta). I. Origin of the group and assessment of the monophyly of its major divisions. *Eur. J. Phycol.* 28: 261-275.

Medlin, L. K., Kooistra, W. H. C. F., Gersonde, R. and Wellbrock, U. 1996a. Evolution of the diatoms (Bacillariophyta): II. Nuclear-encoded small-subunit rRNA sequence comparisons confirm a paraphyletic origin for the centric diatoms. *Mol. Biol. Evol*. 13: 67-75.

Medlin, L. K., Kooistra, W. H. C. F., Gersonde, R. and Wellbrock, U. 1996. Evolution of the diatoms (Bacillariophyta): III. Molecular evidence for the origin of the Thalassiosirales. *Nov. Hedwig. Beih.* 112: 221-234.

Medlin, L. K. and Kaczmarska, I. 2004. Evolution of the diatoms: V. Morphological and cytological support for the major clades and a taxonomic revision. *Phycologia*. 43: 245-270.

Moustafa, A., Beszteri, B., Maier, U. G., Bowler, C., Valentin, K., Bhattacharya, D. 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science*. 324: 1724-1726.

Nakajima, K., Tanaka, A., Matsuda, Y. 2013. SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc. Natl. Acad. Sci. U. S. A.* 110: 1767-1772.

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I. V., Dubchak, I. 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res*. 42(1): D26-31.

Paradis, E., Claude, J. and Strimmer, K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*. 20(2): 289-290.

Parisi, G., Perales, M., Fornasari, M. S., Colaneri, A., Gonza ́lez-Schain, N., Go ́mez-Casati, D., Zimmermann, S., Brennicke, A., Araya, A., Ferry, J. G., Echave, J., Zabaleta, E. 2004. Gamma carbonic anhydrases in plant mitochondria. *Plant Mol Biol*. 55: 193-207.

Price, G. D., Badger, M. R., Woodger, F. J., Long, B. M. 2008. Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): Functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. *J. Exp. Bot*. 59(7): 1441-1461.

Pruesse, E., Peplies, J. and Gl öckner, F. O. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*. 28: 1823-1829.

Raven, J. A. 1994. Carbon fixation and carbon availability in marine phytoplankton. *Photosynth Res*. 39: 259-273.

Raven, J. A., Brown, K., Mackay, M., Beardall, J., Giordano, M., Granum, E., Leegood, R. C., Kilminster, K. and Walker, D. I. 2005. Iron, nitrogen, phosphorus and zinc cycling and consequences for primary productivity in the oceans, In Society for General Microbiology Symposium 65 Micro-organisms and Earth systems: Advances in Geobiology (eds. G M Gadd, K T Semple and H M Lappin-Scott), pp. 247-272. Cambridge University Press, Cambridge.

Raven, J. A., Cockell, C. S., La Rocha, C. L. 2008. The evolution of inorganic carbon concentrating mechanisms in photosynthesis. *Phil Trans Roy Soc B.* 363: 2641-2650.

Raven, J. A., Beardall, J., Giordano, M. and Maberly, S. C. 2011. Algal and aquatic plant carbon concentrating mechanisms in relation to environmental change. *Photosynth Res*. 109: 281-296.

Raven, J. A., Giordano, M., Beardall, J. and Maberly, S. C. 2012. Algal evolution in relation to atmospheric CO_2: carboxylases, carbon-concentrating mechanisms and carbon oxidation cycles. *Phil. Trans. R. Soc. B.* 367: 493-507.

Reinfelder, J. R., Kraepiel, A. M. L., Morel, F. M. M. 2000. Unicellular C_4 photosynthesis in a marine diatom. *Nature*. 407: 996-999.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*. 3: 217–223.

Roberts, K., Granum, E., Leegood, R. C., Raven, J. A. 2007. C_3 and C_4 pathways of phototsynthetic carbon assimilation in marine diatoms are under genetic, not environmental control. *Plant physiol*. 145: 230-235.

Roberts, S. B., Lane, T. W., Morel, F. M. M. 1997. Carbonic anhydrase in the marine diatom *Thalassiosira weissflogii* (Bacillariophyceae). *J Phycol*. 33: 845-850.

Rost, B., Riebesell, U., Burkhardt, S., Sültemeyer, D. 2003. Carbon acquisition of bloom-forming marine phytoplankton. *Limnol Oceanogr* 48: 55-67.

Round, F. E., Crawford, R. M. and Mann, D. G. 1990. The Diatoms: Biology & Morphology of the Genera. Cambridge, UK: Cambridge University Press.

Samukawa, M., Shen. C., Hopkinson, B. M., Matsuda, Y. 2014. Localization of putative carbonic anhydrases in the marine diatom, *Thalassiosira pseudonana*. *Photosynthesis Res*. 121: 235-249.

Simonsen, R. 1972. Ideas for a more natural system of the centric diatoms. *Nova Hedwigia*. 29: 37-54.

Simonsen, R. 1979. The diatom system: Ideas on phylogeny. *Bacillaria*. 2: 9-71.

Sims, P. A., Mann, D. G., Medlin, L. K. 2006. Evolution of the diatoms: insights from fossil, biological and molecular data. *Phycologia*. 45: 361-402.

Sorhannus, U. 2004. Diatom phylogenetics inferred based on direct optimization of nuclear-encoded SSU rRNA sequences. *Cladistics*. 20: 487-497.

Sorhannus, U. 2007. A nuclear-encoded small-subunit ribosomal RNA timescale for diatom evolution. *Mar. Micropaleontol*. 65: 1-12.

Tabita, F. R., Hanson, T. E., Li, H., Satagopan, S., Singh, J. and Chan, S. 2007. Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.* 71: 576.

Tabita, F. R., Hanson, T. E., Satagopan, S., Witte, B. H. and Kreel, N. E. 2008a. Phylogenetic and evolutionary and the functional lessons provided by diverse molecular forms. *Phil. Trans. R. Soc. B.* 363: 2629-2640.

Tabita, F. R., Satagopan, S., Hanson,T. E., Kreel, N. E. and Scott, S. S. 2008b. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.* 59: 1515-1524.

Tachibana, M., Allen, A. E., Kikutani, S., Endo, Y., Bowler, C., Matsuda, Y. 2011. Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Photosynth Res*. 109: 205-221.

Talavera, G., and Castresana, J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*. 56: 564-577.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*. 30: 2725-2729.

Theriot, D. C., Cannone, J. J., Gutell, R. R. and Alverson, A. J. 2009. The limits of nuclear-encoded SSU rDNA for resolving the diatom phylogeny. *Eur. J. Phycol.* 44(3): 277-290.

Theriot, E. C., Ashworth, M., Ruck, E., Nakov, T., and Jansen, R. K. 2010. A preliminary multigene phylogeny of the diatoms (Bacillariophyta): challenges for future research. *Plant Ecology and Evolution*. 143(3): 278-296.

Tréguer, P., Nelson, D. M., Bennekom, A. J., DeMaster, D. J., Leynaert, A., Quéquiner, B. 1995. The silica balance in the world ocean: a reestimate. *Science*. 268: 375-379.

Trimborn, S., Lundholm, N., Thomas, S., Richter, K. U., Krock, B., Hansen, P. J., Rost, B. 2008. Inorganic carbon acquisition in potentially toxic and non-toxic diatoms: the effect of pH-induced changes in seawater carbonate chemistry. *Physiol Plant*. 133: 92-105.

Trimborn, S., Wolf-Gladrow, D., Richter, K-U., Rost, B. 2009. The effect of pCO_2 on carbon acquisition and intracellular assimilation in four marine diatoms. *Journal of Experimental Marine Biology and Ecology.* 376: 26-36.

Young, J. N., Rickaby, R. E. M., Kapralov, M. V. and Filatov, D. A. 2012. Adaptive signals in algal Rubisco reveal a history of ancient atmospheric carbon dioxide. *Phil. Trans. R. Soc. B.* 367: 483-492.

Table 4.1. Diatom strains used in this study.

Species	Abberivation	Shape
Amphora coffeaeformis-CCMP127	*A. coffeaeformis*	raphid pennate
Amphiprora sp	*Amphiprora sp*	raphid pennate
Fragilariopsis kerguelensis-L2_C3	*F. kerguelensis*	raphid pennate
Fragilariopsis cylindrus-CCMP1102	*F. cylindrus*	raphid pennate
Nitzschia punctata-CCMP561	*N. punctata*	raphid pennate
Pseudo_nitzschia fradulenta-WWA7	*P. fradulenta*	raphid pennate
Pseudo_nitzschia australis-10249_10_AB	*P. australis*	raphid pennate
Pseudo_nitzschia multiseries-CLN-47	*P. multiseries*	raphid pennate
Phaeodactylum tricornutum	*P. tricornutum*	raphid pennate
Asterionellopsis glacialis-CCMP134	*A. glacialis*	araphid pennate
Thalassiothrix antarctica-L6_D1	*T. antarctica-L6D1*	araphid pennate
Thalassionema nitzschioides-L26_B	*T. nitzschioides*	araphid pennate
Chaetoceros affinis-CCMP159	*C. neogracilep*	polar centric
Chaetoceros curvisetus	*C. affinis*	polar centric
Chaetoceros debilis-MM31A_1	*C. debilis*	polar centric
Chaetoceros neogracile-CCMP1317	*C. curvisetus*	polar centric
Ditylum brightwellii-GSO103	*D. brightwellii-103*	polar centric
Ditylum brightwellii-GSO104	*D. brightwellii-104*	polar centric
Ditylum brightwellii-GSO105	*D. brightwellii-105*	polar centric
Extubocellulus spinifer-CCMP396	*E. spinifer*	polar centric
Corethron pennatum-L29A3	*C. pennatum*	radial centric
Proboscia alata-PI_D3	*P. alata*	radial centric
Skeletonema dohrnii-SkelB	*S. dohrnii*	radial centric
Skeletonema marinoi-SkelA	*S. marinoi*	radial centric
Skeletonema menzelii-CCMP793	*S. menzelii*	radial centric
Thalassiosira antarctica-CCMP982	*T. antarctica-982*	radial centric
Thalassiosira gravida-GMp14c1	*T. gravida*	radial centric
Thalassiosira miniscula-CCMP1093	*T. miniscula*	radial centric
Thalassiosira oceanica-CCMP1005	*T. oceanica*	radial centric
Thalassiosira pseudonana-CCMP1335	*T. pseudonana*	radial centric
Thalassiosira rotula-CCMP3096	*T. rotula-3096*	radial centric
Thalassiosira rotula-GSO102	*T. rotula-GSO102*	radial centric
Thalassiosira weissflogii-CCMP1010	*T. weissflogii-1010*	radial centric
Thalassiosira weissflogii-CCMP1336	*T. weissflogii-1336*	radial centric

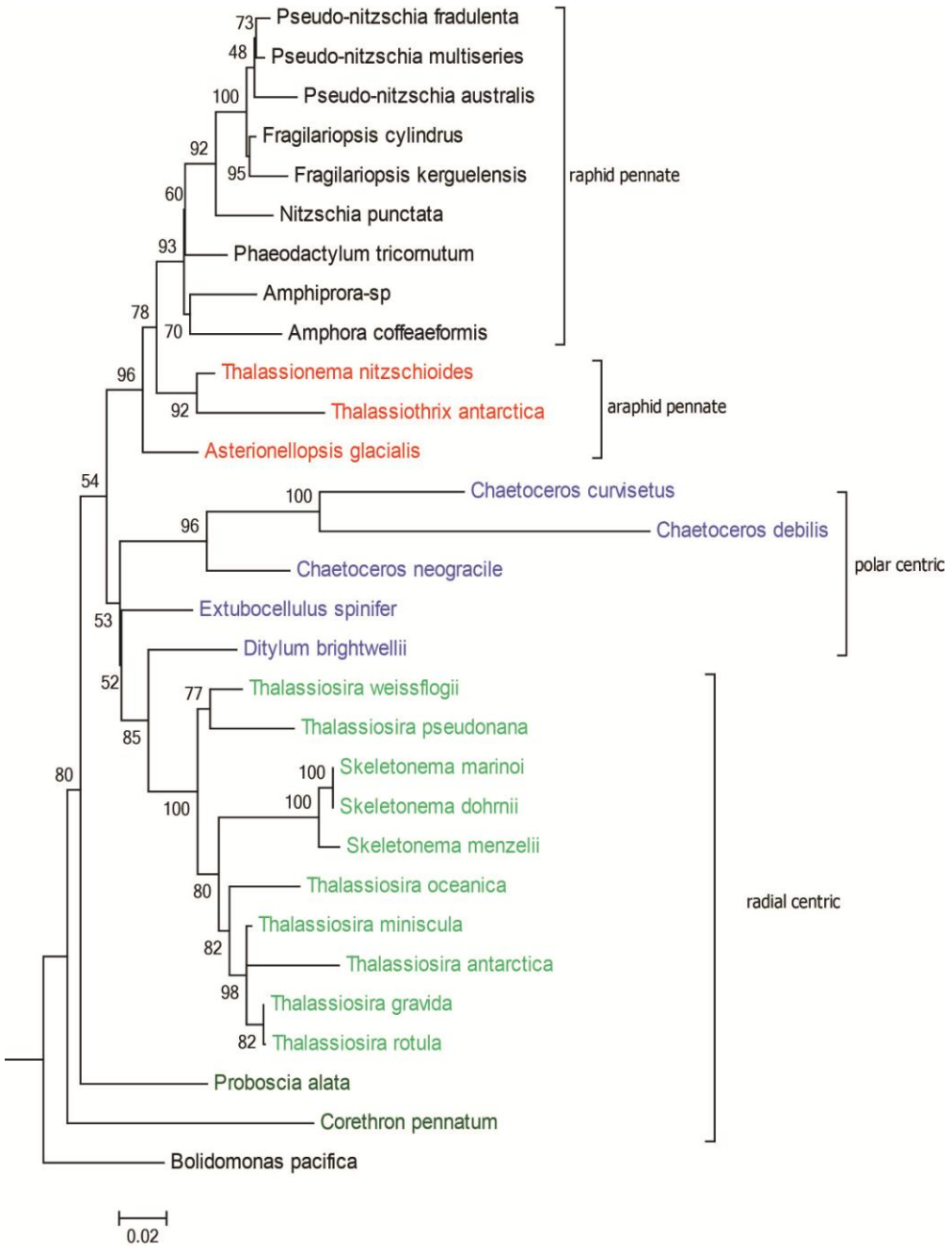
Table 4.2. Clustering result from OrthoMCL analysis. Number represents how many strains out of a total of 34 possess the sequences in that group. Italic font indicates the group contained less than 4 sequences, which were removed when building hierarchical clustering to be consistent with grouping standard of phylogenetic method.

	HCO_3^- transporters	α-CA	γ-CA	δ-CA
group1	28	31	32	24
group2	24	15	32	7
group3	20	14	32	5
group4	19	13	20	5
group5	9	5	9	3
group6	-	5	3	*3*
group7	-	4	*3*	*2*
group8	-	*3*	*1*	*2*
group9	-	*3*	-	*2*
group10	-	*3*	-	*2*
group11	-	*3*	-	-
group12	-	*2*	-	-
group13	-	*2*	-	-
group14	-	*2*	-	-
group15	-	*1*	-	-
group16	-	*1*	-	-
group17	-	*1*	-	-
group18	-	*1*	-	-
total strains	34	34	34	34

Table 4.3. CCM gene grouping by OrthoMCL and maximum parsimony trees.

	Total protein sequences	After manually curated	clustered sequences		clustered groups	
			OrthoMCL	MP tree	OrthoMCL	MP tree
HCO_3^- transporters	154	134	124	108	5	7
α-CAs	245	190	165	50	18	9
γ-CAs	149	138	135	84	8	12
δ-CAs	110	84	72	75	10	7

A


Pseudo-nitzschia fradulenta
Pseudo-nitzschia multiseries
Pseudo-nitzschia australis
Fragilariopsis cylindrus
Fragilariopsis kerguelensis
Nitzschia punctata
Phaeodactylum tricornutum
Amphiprora-sp
Amphora coffeaeformis
raphid pennate
Thalassionema nitzschioides
Thalassiothrix antarctica
araphid pennate
Asterionellopsis glacialis
Chaetoceros curvisetus
Chaetoceros debilis
Chaetoceros neogracile
Extubocellulus spinifer
Ditylum brightwellii
polar centric
Thalassiosira weissflogii
Thalassiosira pseudonana
Skeletonema marinoi
Skeletonema dohrnii
Skeletonema menzelii
Thalassiosira oceanica
Thalassiosira miniscula
Thalassiosira antarctica
Thalassiosira gravida
Thalassiosira rotula
radial centric
Proboscia alata
Corethron pennatum
Bolidomonas pacifica
0.02

B

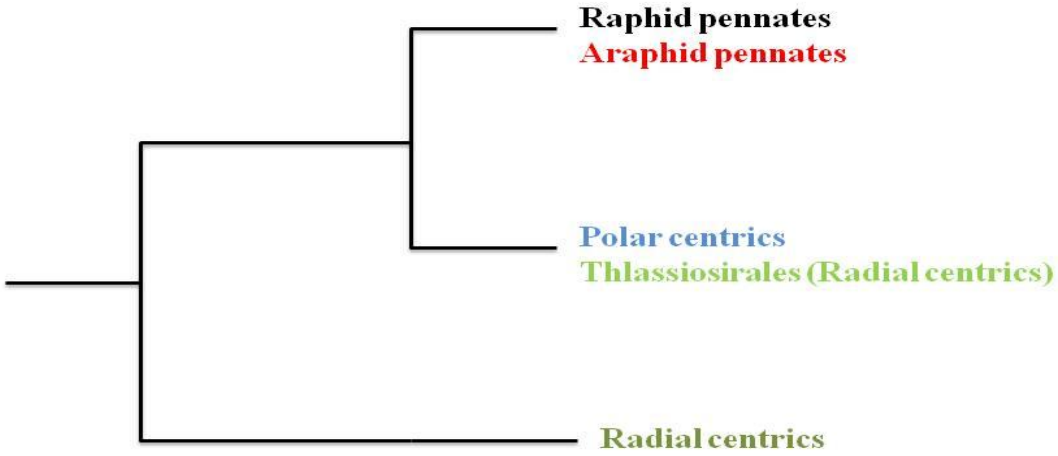
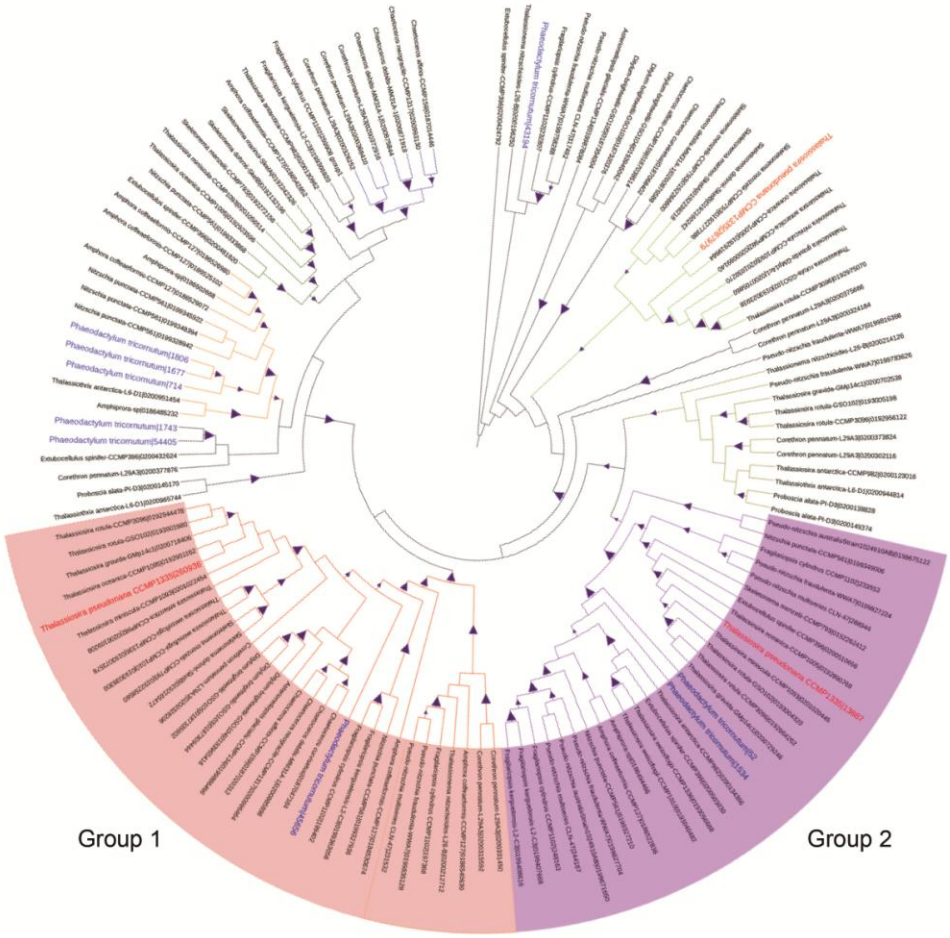


Figure 4.1. A. 18S rDNA phylogenetic tree of diatoms strains listed in Table 4.1. The Maximum likelihood tree was generated by Mega 6 based on 18S rDNA sequences from MMETSP and Genbank. The outer group is *Bolidomonas pacifica* (Genbank ID HQ912557.1). B. Schematic representations of the molecular phylogeny from Medlin and Kaczmarska (2004) showing the relationship between the major lineages within the diatoms. Redrawn according to Kaczmarska et al. (2005). Black: raphid pennates; Red: araphid pennates; Blue: polar centrics; Light green: radial centrics belonging to order Thalassiosirales; Dark green: radial centrics.

A



B

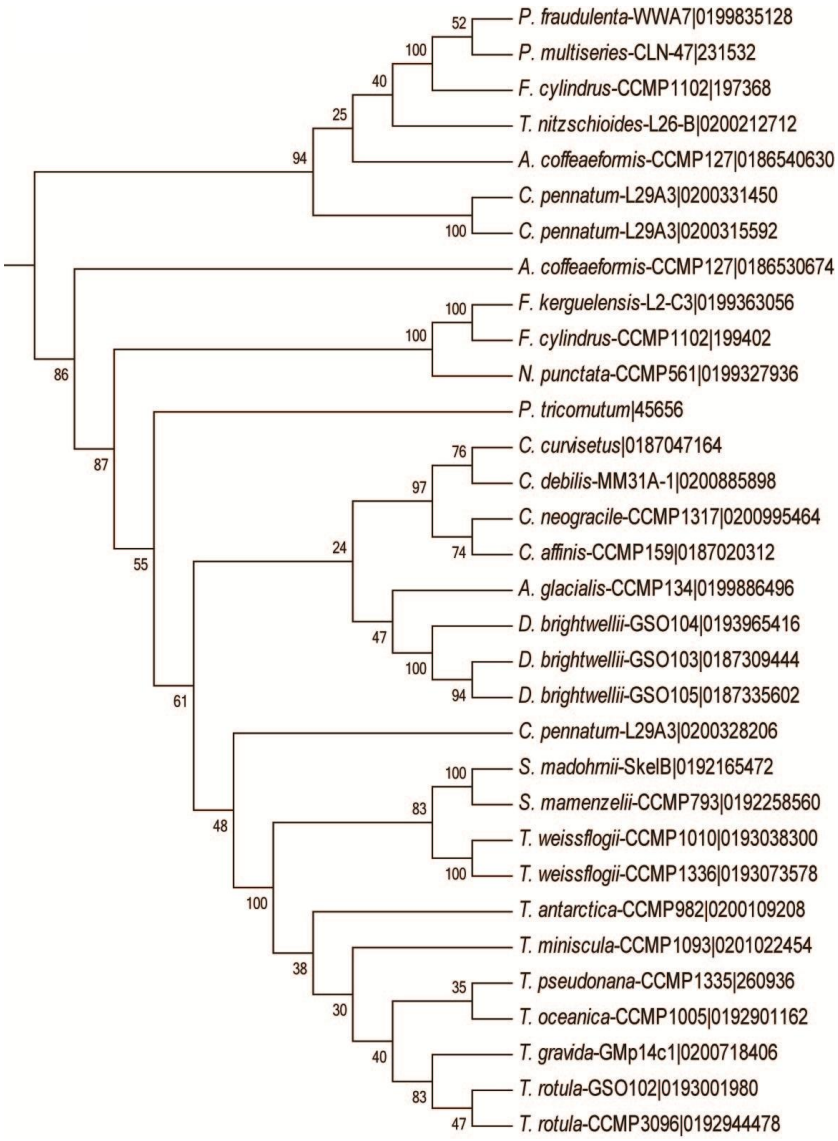


Figure 4.2. A. Maximum parsimony tree of possible bicarbonate transporters. Triangle symbols indicate bootstrap value greater than or equal to 50% and different colors indicate different groups determined using the criteria described in the Methods and Material. The red text indicate HCO_3^- transporters from *T. pseudonana* and blue text indicates sequences from *P. tricornutum*. The first two large groups contained sequences from most strains. Group 1 (red clade in the circular tree) has sequences from 28 strains is shown in panel B and group 2 (purple clade in the circular tree) has sequences from 20 strains; B. Subtrees Group 1 (red clade in the circular tree) contains sequences from 28 strains.

A


Group 2
Group 3
Group 1
Group 4
Thalassiosira pseudonana-CCMP1335|814
Thalassiosira pseudonana-CCMP1335|34094
Thalassiosira pseudonana-CCMP1335|233
Thalassiosira pseudonana-CCMP1335|34125

B

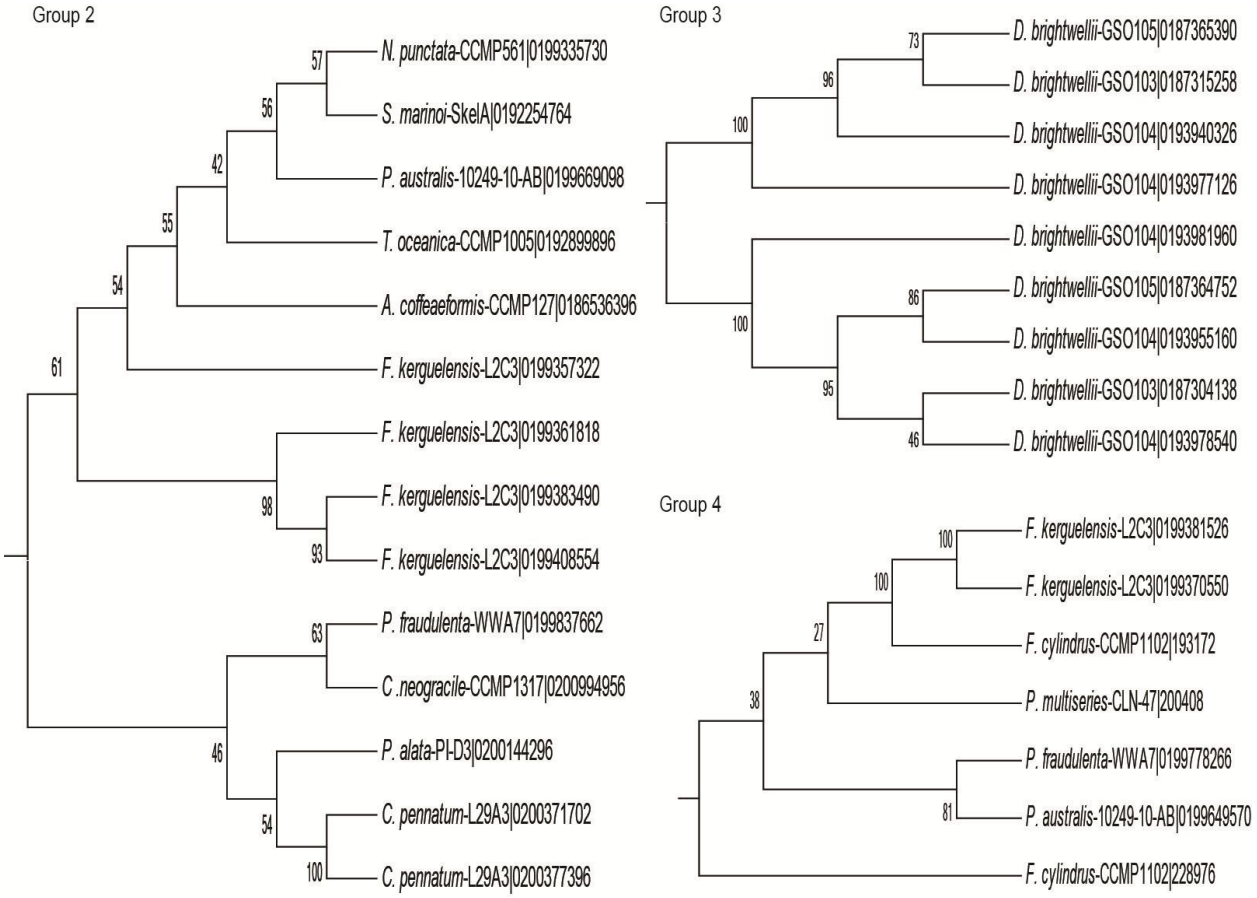
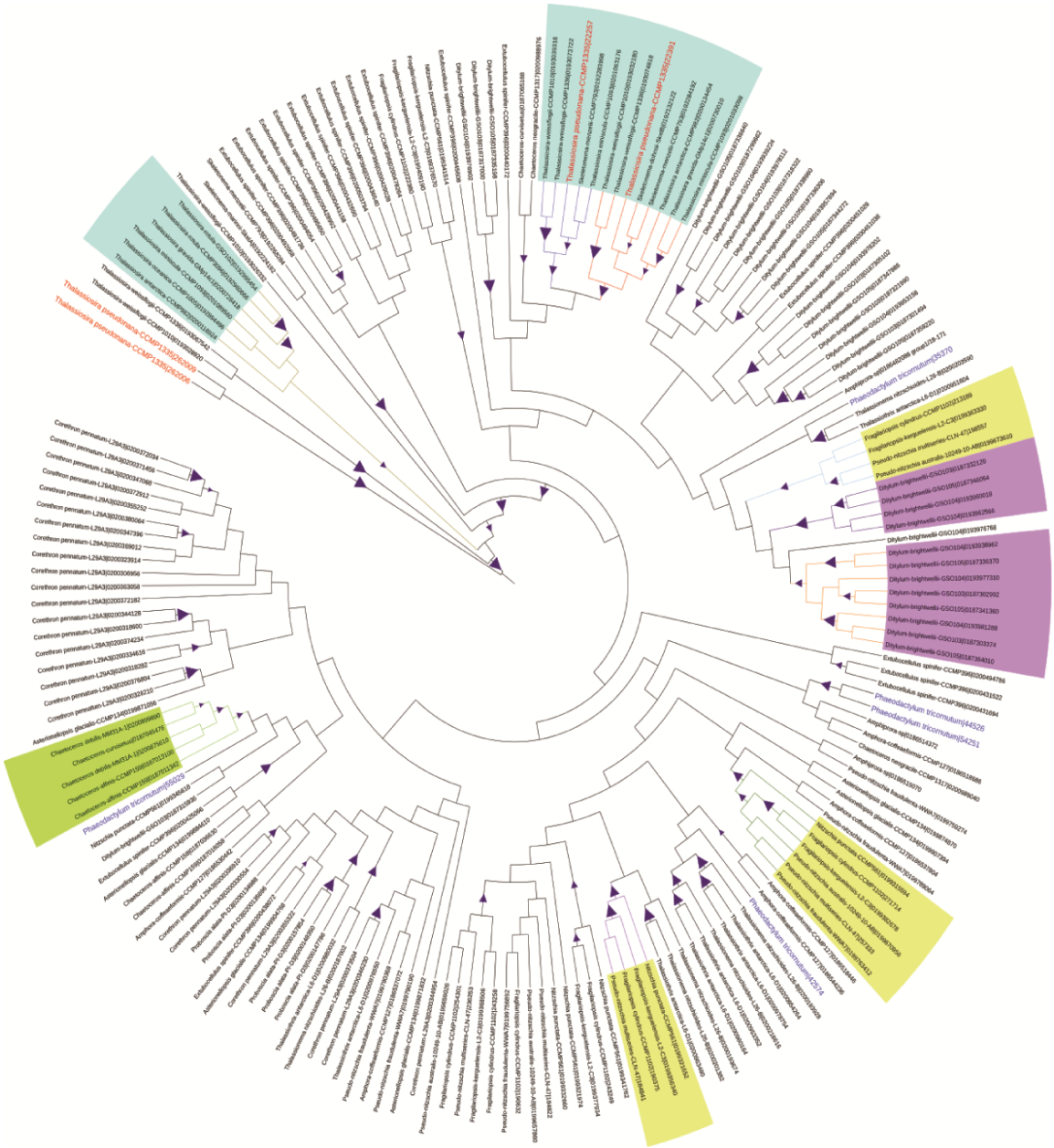


Figure 4.3. A. Maximum parsimony tree of possible δ-CAs. Triangle symbols indicate bootstrap value greater than or equal to 50% and different colors indicate different groups determined using the criteria described in the Methods. The red text indicates δ-CAs from *T. pseudonana*. Four major groups were identified. Group 1 (blue clade in the circular tree) contained sequences from *Skeletonema* and *Thalassiosira* species, Group 2 (red clade) contained sequences from multiple strains, Group 3 (purple) possess genes from *Ditylum brightwellii* strains and sequences forming Group 4 (yellow) came from *Pseudo-nitzschia* and *Fragilariopsis* species. B. Subtrees from maximum parsimony tree of possible δ-CAs: Group 2 (red clade), Group 3 (green), and Group 4 (yellow).

A



B

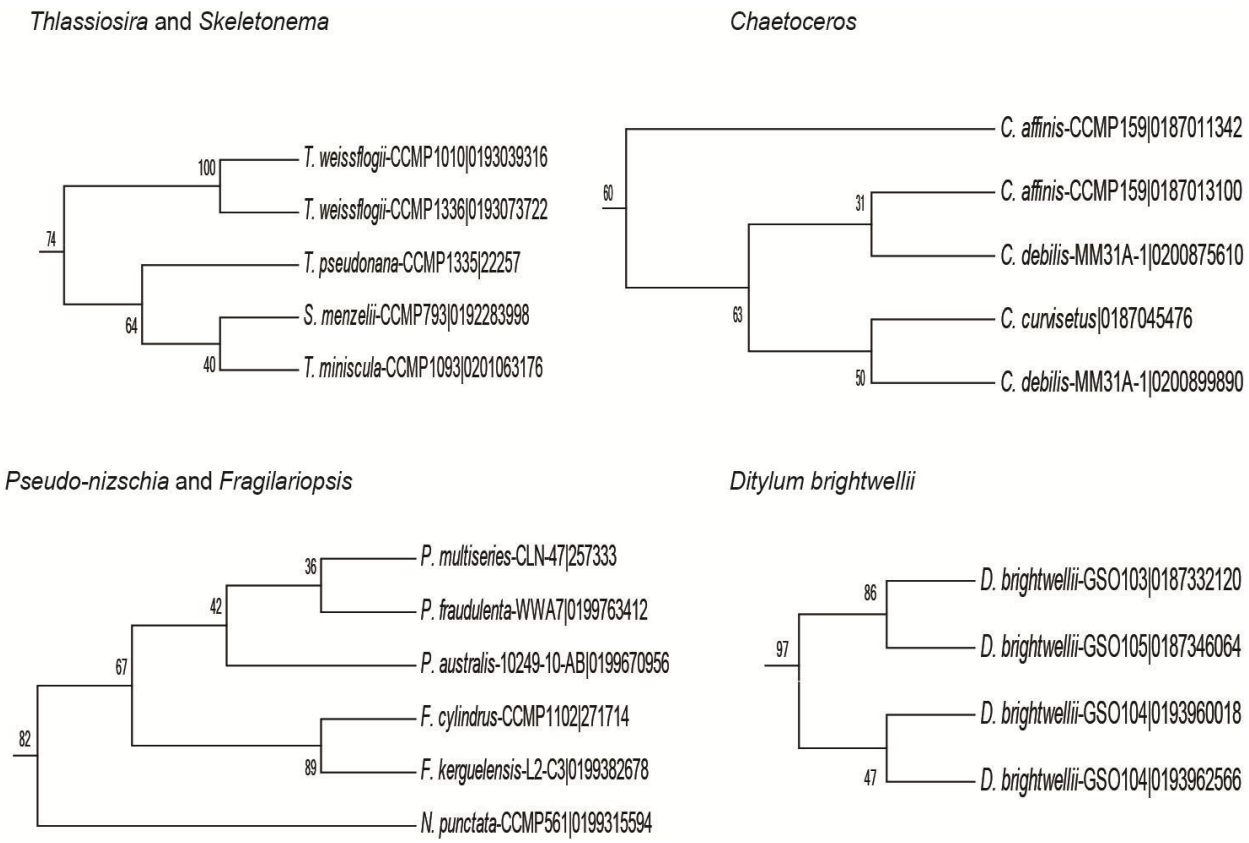
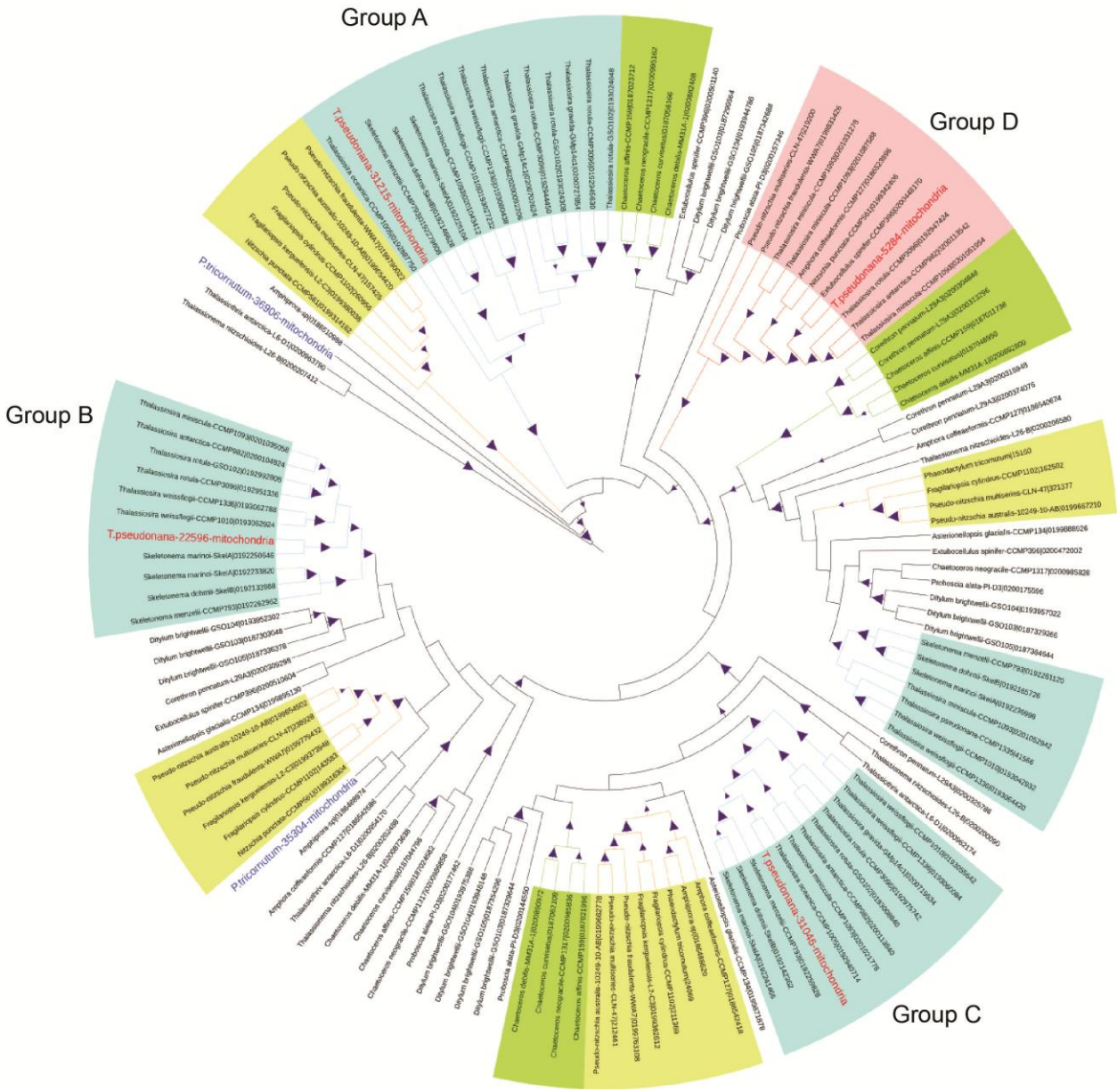


Figure 4.4. A. Maximum parsimony tree of possible α-CAs. Triangle symbols indicate bootstrap value greater than or equal to 50% and different colors indicate different groups determined using the criteria described in the Methods. The red text indicate HCO_3^- transporters from *T. pseudonana* and blue text indicates sequences from *P. tricornutum*. (Blue: *Skeletonema* and *Thalassiosira* species; Yellow: *Pseudo-nitzschia* and *Fragilariopsis;* Green: *Chaetoceros* species; Purple: *Ditylum brightwellii* strains). B. Sub-trees from Maximum Parsimony Tree of possible α-CAs. These groups containing sequences from *Skeletonema* and *Thalassiosira* species, *Ditylum brightwellii* strains, *Pseudo-nitzschia* and *Fragilariopsis* species and *Chaetoceros* species.

A


Group A
Group B
Group C
Group D
T.pseudonana-31215-mitochondria
P.tricornutum-36906-mitochondria
T.pseudonana-22596-mitochondria
P.tricornutum-35304-mitochondria
T.pseudonana-5284-mitochondria
T.pseudonana-31046-mitochondria

B

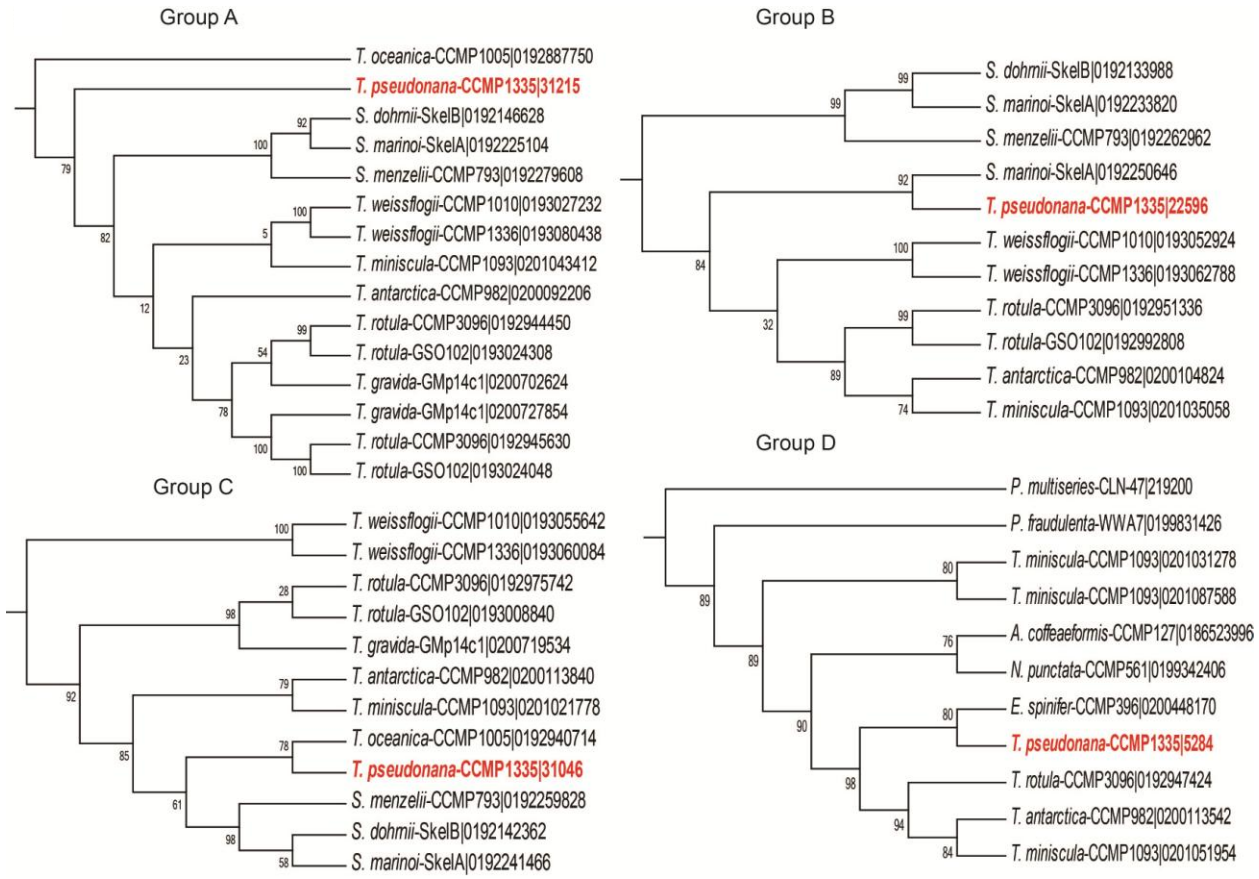


Figure 4.5. A. Maximum parsimony tree of possible γ-CAs. Triangle symbols indicate bootstrap value greater than or equal to 50% and different colors indicate different groups determined using the criteria described in the Methods. The major clades are colored and indicate groups containing sequences from certain species (Blue: *Skeletonema* and *Thalassiosira* species; Yellow: *Pseudo-nitzschia* and *Fragilariopsis;* Green: *Chaetoceros* species; Red: both centrics and pennates). γ-CAs from *T. pseudonana* and *P. tricornutum* which have been localized to mitochondria (Tachibana et al., 2011; Samukawa et al., 2014) were shown as red and blue text respectively in the circular tree. B. The four sub-trees contained γ-CAs from *T. pseudonana* that have been shown or are predicted to be localized to the mitochondrion. (Tp5284 in groupD is predicted to be localized in mitochondrion)

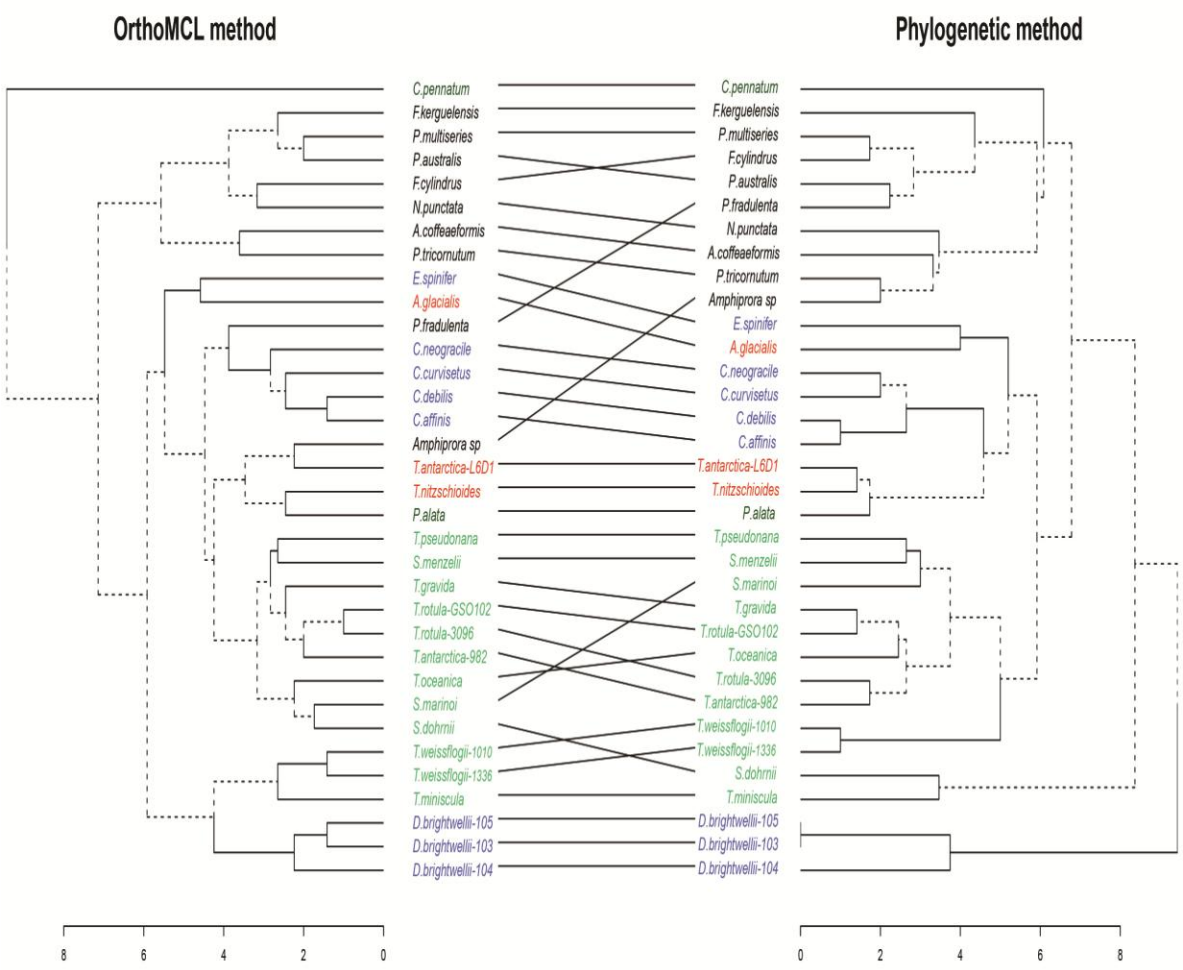
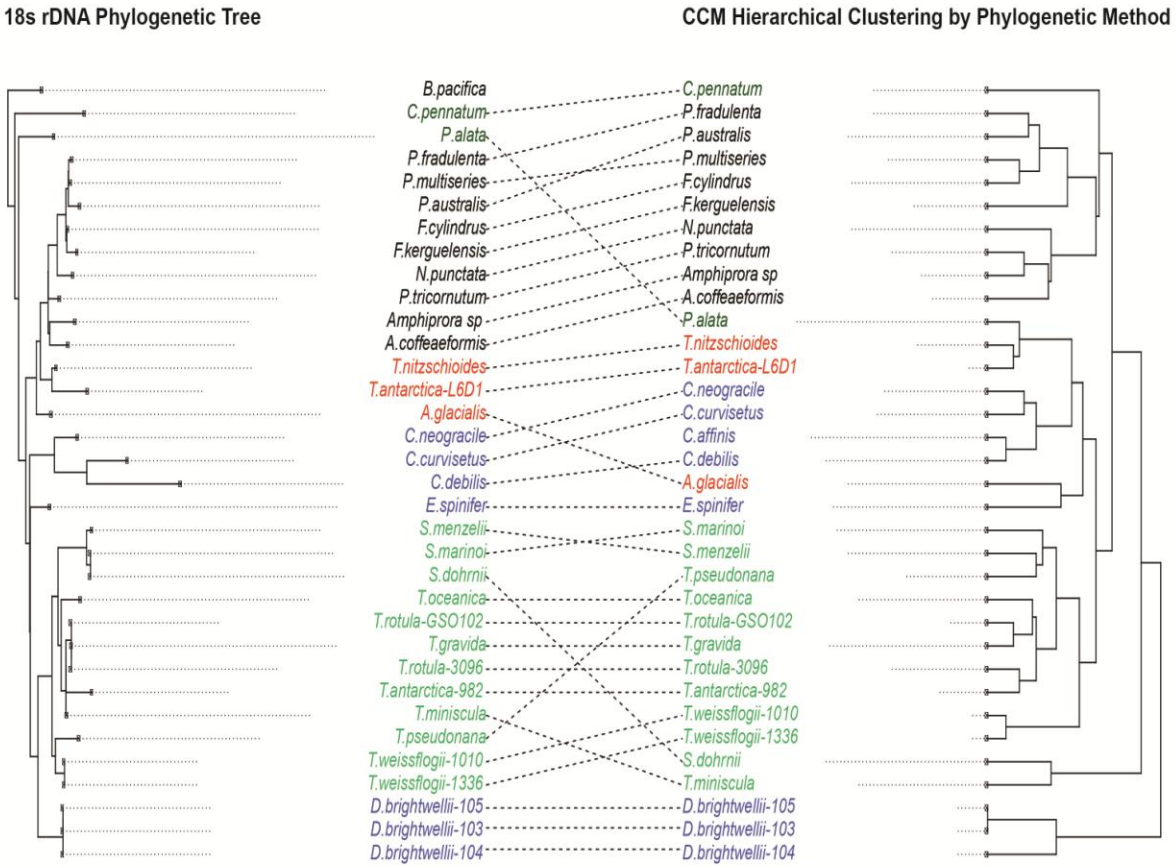


Figure 4.6. Comparison of two hierarchical clusterings of 34 diatom strains in terms of their CCM gene content as grouped by OrthoMCL and protein phylogeny. Black: raphid pennates; Red: araphid pennates; Blue: polar centrics; Light green: radial centrics belonging to order Thalassiosirales); Dark green: radial centrics.

A


18s rDNA Phylogenetic Tree
CCM Hierarchical Clustering by Phylogenetic Method
B.pacifica
C.pennatum
P.alata
P.fradulenta
P.multiseries
P.australis
F.cylindrus
F.kerguelensis
N.punctata
P.tricornutum
Amphiprora sp
A.coffeaeformis
T.nitzschioides
T.antarctica-L6D1
A.glacialis
C.neogracile
C.curvisetus
C.debilis
E.spinifer
S.menzelii
S.marinoi
S.dohrnii
T.oceanica
T.rotula-GSO102
T.gravida
T.rotula-3096
T.antarctica-982
T.miniscula
T.pseudonana
T.weissflogii-1010
T.weissflogii-1336
D.brightwellii-105
D.brightwellii-103
D.brightwellii-104
C.pennatum
P.fradulenta
P.australis
P.multiseries
F.cylindrus
F.kerguelensis
N.punctata
P.tricornutum
Amphiprora sp
A.coffeaeformis
P.alata
T.nitzschioides
T.antarctica-L6D1
C.neogracile
C.curvisetus
C.affinis
C.debilis
A.glacialis
E.spinifer
S.marinoi
S.menzelii
T.pseudonana
T.oceanica
T.rotula-GSO102
T.gravida
T.rotula-3096
T.antarctica-982
T.weissflogii-1010
T.weissflogii-1336
S.dohrnii
T.miniscula
D.brightwellii-105
D.brightwellii-103
D.brightwellii-104

B

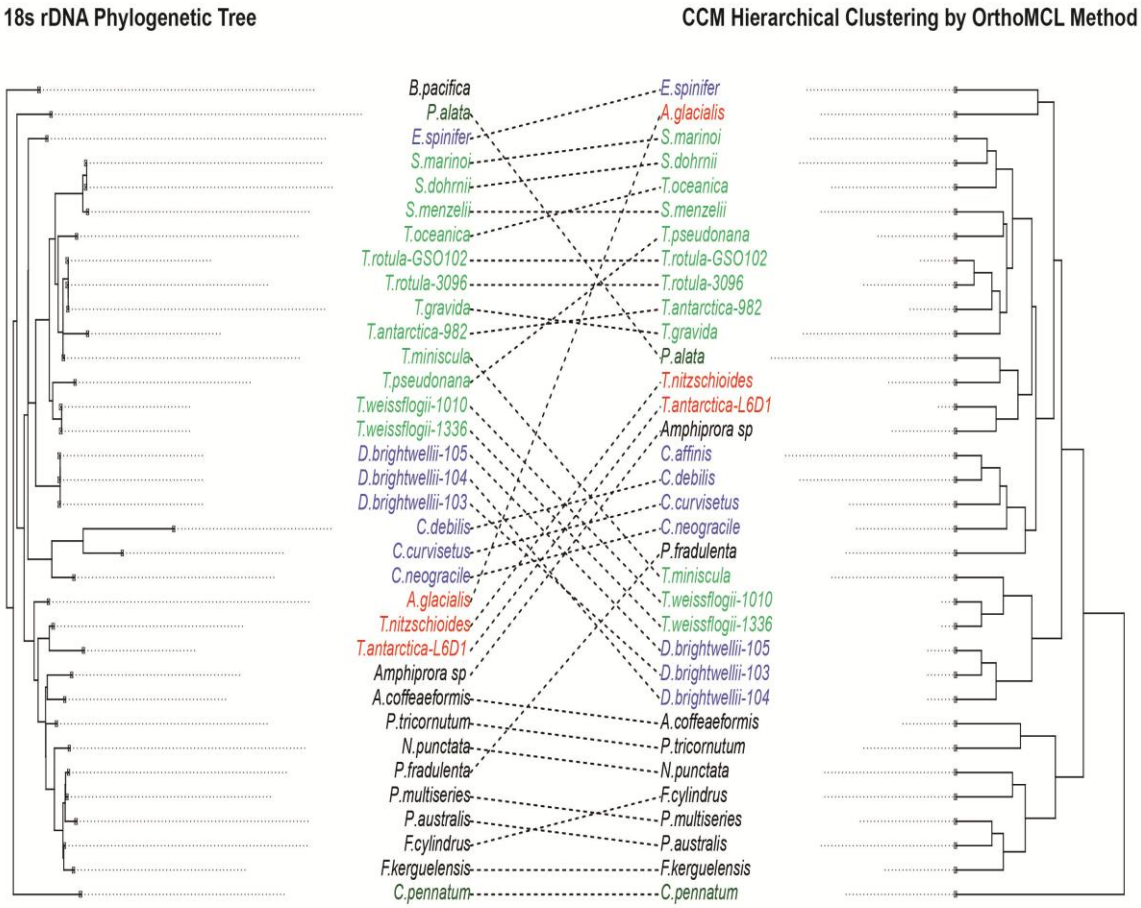


Figure 4.7. A. Comparison of diatom 18S rDNA phylogenetic tree and the dendrogram of CCM genes grouping by protein phylogeny method. B. Comparison of diatom 18S rDNA phylogenetic tree and the dendrogram of CCM genes grouping by OrthoMCL method. Black: raphid pennates; Red: araphid pennates; Blue: polar centrics; Light green: radial centrics belonging to order Thalassiosirales; Dark green: radial centrics.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

Anthropogenic CO_2 release is resulting in a series of environment problems including ocean acidification and global warming (Caldeira and Wickett, 2003; Broecker, 1975). Changes in seawater pH, inorganic carbon chemistry, nutrients, and light caused by ocean acidification and global warming are expected to affect the phytoplankton in different ways (Wolf-Gladrow et al., 1999; Rost et al., 2008). Understanding the net effects of these coupled environmental changes is a difficult but important task as phytoplankton play a major role in oceanic carbon sequestration, thus modulating the global carbon cycle (Reinfelder, 2011; Flynn et al., 2012). As one of the most ecologically significant primary producers in the ocean, marine planktonic diatoms are responsible for up to 20% of global carbon fixation (Tréguer et al., 1995; Nelson et al., 1995; Field et al., 1998; Mann, 1999). Diatoms activily take up HCO_3^- and CO_2 and possess CO_2 concentrating mechanisms (CCMs) to accumulate inorganic carbon (Ci) from seawater, which are essential for their high rates of primary production (Trimborn et al., 2008; Matsuda et al., 2011). The CCM is likely to be one of the primary systems affected by ocean acidification as it does not need to work as hard at high CO_2. Therefore, it is important to understand the CCM in diatoms, since this system in part mediates the response of phytoplankton to anthropogenic perturbations of the global carbon cycle.

The work in this dissertation explores CCMs in diatoms from genetic, physiological, and evolutionary perspectives. CCMs in diatoms are primarily composed of active Ci uptake systems to transport HCO_3^-, carbonic anhydrases to assist accumulation of HCO_3^- and to convert

accumulated HCO_3^- to CO_2, and a confined compartment where RubisCO aggregates to limit the extent of the zone of CO_2 elevation. The locations of CCM components i.e. CAs and HCO_3^- transporters are critical to their functional roles, so in the second chapter I tried to localize certain CA and HCO_3^- transporters in model diatom *Thalassiosira pseudonana* using gene transformation techniques and fluorescence microscopy. A homolog of TWCA in *T. pseudonana* was localized and it formed discrete clusters associated with the chloroplast, which indicated it most likely localized within the periplastidal compartment. This is so far the only CA that has been observed within *T. pseudonana* chloroplast membranes, though there are still several CAs that have yet to be localized (Samukawa et al., 2014). Compared with five α-CAs localized to the chloroplast membrane system in *P. tricornutum*, this illustrates the diversity of CCMs in these two model diatoms. I also localized three putative HCO_3^- transporters, one to the outer membrane and two to the chloroplast membranes in *T. pseudonana*. Although the activity of these HCO_3^- transporter have not been confirmed yet, it is reasonable to assume that they may work to provide a route for HCO_3^- uptake from the external environment and transport into the chloroplast.

In addition to sub-cellular localization, I also assessed the physiological function of the overexpressed CAs and HCO_3^- transporters. For overexpressed CAs, CA activity was measured with ^{18}O exchange assay via Membrane Inlet Mass Spectrometry (Tu et al., 1978; Hopkinson et al., 2011) and one mutant line showed marginally higher activity. For overexpressed HCO_3^- transporters, photosynthetic rates and HCO_3^- uptake rates were measured using a MIMS-based Photosynthesis vs. Ci method (Badger et al., 1994). A putative outer membrane HCO_3^- transporter (BD714) in *P. tricornutum* was overexpressed and both mutant strains were physiologically characterized, showing higher net photosynthetic rates and HCO_3^- uptake rates

compared with wild types at low Ci concentrations. The active function in both strains overexpressing BD714 indicated that the transporter does indeed transport HCO_3^-. This finding, together with previous findings showing that another SLC4 transporter in *P. tricornutum* transports HCO_3^- (Nakajima et al., 2013), confirmed that the mammalian type SLC4 HCO_3^- transporters found in *P. tricornutum* are actual HCO_3^- transporters. My work on genetic and physiological characterization of the HCO_3^- transporters and CAs facilitates building a complete picture of the CCM in model diatoms.

Chapter 2 demonstrates the different locations and activities of CCM genes in *P. tricornutum* and *T. pseudonana,* suggesting a diversity of CCMs in the model diatoms. One of those differences is that extracellular CAs (eCAs) exist in *T. pseudonana* but are absent in *P. tricornutum*. As a component of CCM, one possible role of surface-associated eCAs is to catalyze conversion of HCO_3^- to CO_2 at the surface of phytoplankton cells and the CO_2 generated can then be taken up for photosynthesis. Since larger cells are more susceptible to diffusion limitation, I hypothesized that eCA activity would increase with cell size to support greater CO_2 demand by larger cells and that eCA activity would be of greater importance to photosynthesis in larger cells with the need of higher inorganic carbon supply. eCA activity in six centric diatom species spanning nearly the full range of cell sizes for centric diatoms was quantified by MIMS using a recently developed method (Hopkinson et al., 2013). The result indicated that eCA activity did increase with cell radius by a size-scaling exponent of 2.6 ± 0.3. This rapid increase keeps the absolute CO_2 concentration gradient between bulk seawater and the cell surface very low, allowing high rates of CO_2 uptake even for large diatoms. The effect of eCA activity on photosynthesis was assessed by comparing the carbon fixation rates, with and without an eCA inhibitor, determined from the incorporation of ^{14}C-DIC into biomass. There

was no overall relationship between the extent to which loss of eCA-activity inhibited photosynthesis and cell size. But we demonstrated that while photosynthesis in the smallest diatoms (<4 μm radius) was only affected by eCA inhibition under very low CO_2 concentrations, photosynthesis in some larger diatoms was affected even at typical seawater CO_2 concentrations, indicating that eCA is more important for larger diatoms. The ubiquity of eCA in centric diatoms suggests that eCA is especially important for CO_2 uptake in this group.

The diversity of CCM architecture in two model diatoms (Tachibana et al., 2011; Samukawa et al., 2014; Nakajima et al., 2013) inspired me to explore the CCM in other diatoms, to determine if they all possess similar CCM genes and whether the CCMs of the two model diatoms are representative of diatom CCMs in general. In chapter 4, I used bioinformatic methodology to screen for potential CCM genes in genomes and transcriptomes from a total of 34 diatom strains. For each CCM component (HCO_3^- transporters and three types of CAs), the sequences identified were clustered into groups of related sequences using two methods: OthoMCL and phylogenetic analysis. The grouping results indicated that these diatoms seem to possess similar HCO_3^- transporters, but have different carbonic anhydrases. Taking all the CCM components into consideration, there is obviously much variety among the diatoms and it is hard to consider the CCMs of *T. pseudonana* and *P. tricornutum* as representatives of diatoms in general. The relationships among diatom CCMs were examined using hierarchical clusterings based on CCM gene content as determined using OrthoMCL and phylogenetic groupings of sequences. The two clusterings are generally in agreement, and both show certain consistent correspondences between CCM structure and diatom morphology: raphid pennates form a coherent group and polar centrics and the Thalassiosirales generally cluster together, but the radial centrics (*Proboscia alata* PI_D3) and araphid pennates are mixed in haphazardly with the

centrics cluster. Comparison of hierarchical clusterings of diatoms based on their CCM components with a diatom 18S rDNA phylogenetic tree indicated the trees are largely congruent, despite some discrepancies in the groupings and difference in the branching structure. This suggests that CCMs appear to reflect diatom phylogeny, at least at finer taxonomic resolution (i.e. genera and strains).

Overall, this dissertation characterized the genetics, physiology, and diversity of CCMs in marine diatoms, using genetic transformation techniques, Membrane Inlet Mass Spectrometry and bioinformatics analysis. Though the CCMs in two model diatoms, *T. pseudonana* and *P. tricornutum*, have been relatively well studied, much work remains for future research. First of all, though thirteen candidate CA genes have been identified in *T. pseudonana* (Samukawa et al., 2014), some CAs still have not been localized and the activity and role of these CAs in the CCM needs to be further investigated. Most notably the role of the problematic α-CA in the chloroplast stroma needs to be determined, as at first glance it would appear to prevent accumulation of HCO_3^- in the stroma. Beyond CAs, more work needs to be done on characterization of HCO_3^- transporters in diatoms. Nakajima et al. (2013) showed that one of the SLC4-type HCO_3^- transporters in *P. tricornutum* was localized to the plasma membrane, and did function as a HCO_3^- transporter. Putative HCO_3^- transporters have been identified in both *T. pseudonana* and *P. tricornutum* based on sequences analysis (Kroth et al., 2008; Nakajima et al., 2013), but few of these have been localized or physiologically characterized. My work on localization of CA and HCO_3^- transporters in *T. pseudonana* and physiological assessment of HCO_3^- transporter activity in *P. tricornutum* contributes to builiding a complete picture of CCM in these two model diatoms, but further investigations on the role and activity of CCM components are still needed.

Another promising direction is to extend the study of eCA conducted in this dissertation. As discussed in chapter 3, eCA is ubiquitous in centric marine diatoms and this distribution suggests CCM physiology has unique characteristics in this group. Considering the complicated evolutionary history of diatoms, including multiple endosymbiotic events, the possible origin and evolution of eCAs among centric diatoms may shed some light on CCM evolution and diatom phylogeny. To fully understand the function of eCA and its origins, a more detailed characterization of CCM physiology in those diatoms that possess eCAs is required.

Chapter 4 explored the extent of CCM diversity among diatoms and suggested some CCM genes may diversify within diatom lineages. The dataset is composed of mostly transcriptomes from marine diatoms. As the genomes of more eukaryotes get sequenced, additional genetic data will be collected to support more accurate and robust phylogenetic analysis and more thorough evaluation of CCM genes. On the other hand, this chapter covered analysis of α-CAs, γ-CAs, and δ-CAs, future study should continue to work on β, ε, and ζ type CAs to explore the diversity of CCM components, which requires further identification and characterization of these CAs.

References

Badger, M. R., and Price, G. D. 1994. The role of carbonic anhydrase in photosynthesis. *Ann Rev Plant Physiol Plant Mol Biol*. 45: 369-392.

Broecker, W. S. 1975. Climatic Change- Are we on brink of a pronounced global warming? *Science*. 189: 460- 463.

Caldeira, K. and Wickett, M. E. 2003. Anthropogenic carbon and ocean pH. *Nature*. 425: 365.

Falkowski, P., Scholes, R. J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Hoegberg, P., Linder, S., Mackenzie, F. T., Moore III, B., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., Steffen, W. 2000. The global carbon cycle: a test of our knowledge of Earth as a system. *Science*. 290: 291-296.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., Falkowski, P. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. 281: 237-240.

Flynn, K. J., Blackford, J. C., Baird, M. E., Raven, J. A., Clark, D. R., Beardall, J., Brownlee, C., Fabian, H., Wheeler, G. L. 2012. Changes in pH at the exterior surface of plankton with ocean acidification. *Nat. Clim. Change*. 2: 510-513.

Hopkinson, B. M., Dupont, C. L., Aleen, A. E., Morel, F.M.M. 2011. Efficiency of the CO_2-concentrating mechanism of diatoms. *Proc. Nat. Acad. Sci*. 108: 3830-3837.

Hopkinson, B. M., Meile, C., and Shen, C. 2013. Quantification of extracellular carbonic anhydrase activity in two marine diatoms and investigation of its role. *Plant Physiol*. 162: 1142-1152.

Kroth, P. G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M. S., Stanley, M. S., Kaplan, A., Caron, L., Weber, T., Maheswari, U., Armbrust, E. V., Bowler, C. 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS One*. 3: e1426.

Mann, D. G. 1999. The species concept in diatoms. *Phycologia*. 38: 437-495.

Matsuda, Y., Nakajima, K., Tachibana, M. 2011. Recent progresses on the genetic basis of the regulation of CO_2 acquisition systems in response to CO_2 concentration. *Photosynth Res*. 109: 191-203.

Nakajima, K., Tanaka, A., Matsuda, Y. 2013. SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc. Natl. Acad. Sci. U. S. A.* 110: 1767-1772.

Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A. & Quéguiner, B. 1995. Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biochemical Cycles*. 9: 359-372.

Reinfelder, J. R. 2011. Carbon concentrating mechanisms in eukaryotic marine phytoplankton. *Ann Rev Mar Sci*. 3: 291-315.

Rost, B., Zondervan, I., Wolf-Gladrow, D. 2008. Sensitivity of the phytoplankton to future changes in ocean carbonate chemistry: current knowledge, contradictions and research directions. *Mar. Ecol. Prog. Ser.* 373: 227-237.

Samukawa, M., Shen, C., Hopkinson, B. M., Matsuda, Y. 2014. Localization of putative carbonic anhydrases in the marine diatom, *Thalassiosira pseudonana*. *Photosynth Resh*. 121: 235-249.

Tachibana, M., Allen, A. E., Kikutani, S., Endo, Y., Bowler, C., Matsuda, Y. 2011. Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Photosynth Res*. 109: 205-221.

Tréguer, P., Nelson, D. M., Bennekom, A. J., DeMaster, D. J., Leynaert, A., Quéquiner, B. 1995. The silica balance in the world ocean: a reestimate. *Science*. 268: 375-379.

Trimborn, S., Lundholm, N., Thomas, S., Richter, K. U., Krock, B., Hansen, P. J., Rost, B. 2008. Inorganic carbon acquisition in potentially toxic and non-toxic diatoms: the effect of pH-induced changes in seawater carbonate chemistry. *Physiol Plant*. 133: 92-105.

Tu, C., Wynns, G. C., McMurray, R. E., and Silverman, D. N. 1978. CO_2 kinetics in red cell suspensions measured by ^{18}O exchange. *J. Biol. Chem*. 253: 8178-8184.

Van, K. and Spalding, M. H. 1999. Periplasmic carbonic anhydrase structural gene (Cah1) mutant in *Chlamydomonas reinhardtii*. *Plant Physiol*. 120: 757-764.

Wolf-Gladrow, D. A., Riebesell, U., Burkhardt, S., Bijma, J. 1999. Direct effects of CO_2 concentration on growth and isotopic composition of marine plankton. *Tellus B Chem Phys Meterol*. 51: 461-476.