2010-07-01

New Insights into the Diversity, Distribution and Ecophysiology of Marine Picoeukaryotes

Marie Laure Cuvelier
University of Miami, mcuvelier@rsmas.miami.edu

UNIVERSITY OF MIAMI

NEW INSIGHTS INTO THE DIVERSITY, DISTRIBUTION AND ECOPHYSIOLOGY OF MARINE PICOEUKARYOTES

By

Marie L. Cuvelier

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

June 2010

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

NEW INSIGHTS INTO THE DIVERSITY, DISTRIBUTION AND ECOPHYSIOLOGY OF MARINE PICOEUKARYOTES

Marie L. Cuvelier

Approved:

_____
Alexandra Z. Worden, Ph.D.
Scientist
Monterey Bay Aquarium Research Institute
Moss Landing, California

_____
Terri A. Scandura, Ph.D.
Dean of the Graduate School

_____
Dennis A. Hansell, Ph.D.
Professor
Marine and Atmospheric Chemistry

_____
Gary L. Hitchcock, Ph.D.
Associate Professor and Co-Chair
Marine Biology and Fisheries

_____
Hervé Moreau, Ph.D.
Research Director
Centre National de la Recherche Scientifique
Observatoire Océanologique de Banyuls-sur-Mer
Banyuls-sur-Mer, France

_____
Marjorie F. Oleksiak, Ph.D.
Assistant Professor
Marine Biology and Fisheries

CUVELIER, MARIE (Ph.D., Biology)
New Insights into the Diversity, Distribution and Ecophysiology of Marine Picoeukaryotes (June 2010)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Alexandra Z. Worden.
No. of pages in text. (178)

Marine microbes are an essential component of global biogeochemical cycles. In oligotrophic marine surface waters, the phytoplankton, phototrophic, single-celled (on occasion, colonial) organisms, is often dominated by the picoplankton (cells <2 μm in size), which constitute the base of the marine food chain. The picophytoplankton is composed of three main groups of organisms: two genera of cyanobacteria, *Prochlorococcus* and *Synechococcus,* and a third group, the picoeukaryotes. Even though numerically less abundant than cyanobacteria, picoeukaryotes can contribute significantly to biomass and primary production in this size fraction. Furthermore, picoeukaryotes are a diverse group but this diversity is still underexplored and their ecological roles and physiology is poorly understood. Here uncultured protists are investigated using 18S rRNA gene clone libraries, phylogenetic analyses, specific fluorescence *in situ* hybridization (FISH) probes and other methods in tropical and subtropical waters. Gene sequences comprising a unique eukaryotic lineage, biliphytes, were identified in most samples, whether from high (30 ^{o}C) or low (5 ^{o}C) temperature waters. Sequences within this uncultured group have previously been retrieved from mid and high latitudes. Phycobilin-like fluorescence associated with biliphyte-specific FISH probed cells indicated they may be photosynthetic. Furthermore, the data indicated biliphytes are nanoplanktonic in size, averaging between 3.0 and 4.1 μm. Using the 18S rRNA gene,

sequences belonging to a broadly distributed but uncultivated pico-prymnesiophytes were retrieved. We investigated the ecological importance of these natural pico-prymnesiophyte populations and field experiments showed that they could grow rapidly and contributed measurably to primary production. They also appear to form a large portion of global picophytoplankton biomass, with differing contributions in five biogeographical provinces, from tropical to high latitudes. Finally, the physiology of the picoeukaryote *Micromonas* was studied under a shift from medium to high light and UV radiation. Results showed that the growth of these photosynthetic cells was synchronized with the light: dark period. Forward angle side scatter and red autofluorescence from chlorophyll increased throughout the light period and decreased during the dark period. This is consistent with cell division occurring at the beginning of the dark period. Additionally, genes proposed to have roles in photoprotection were up-regulated under high light and UV, but not in controls.

ACKNOWLEDGMENTS

This research and dissertation would not have been possible without the help, guidance, encouragement, support, friendship and love of so many people and I am infinitely grateful to all of them.

First, I offer my sincere thanks to my committee members. My advisor, Alex Worden, has given me great scientific guidance and support. Alex, thank you for taking a chance with me, for believing in me and for the all the opportunities you have given me. Other members of my committee, Margie Oleksiak, Dennis Hansell, Gary Hitchcock and Hervé Moreau, have also provided me with great advice and encouragement and I consider myself fortunate to be guided by such a remarkable group of scientists.

At sea, the Captain, Shawn Lake, and crew of the *F.G. Walton Smith* and Captain, Diego Mello, and crew of the *Oceanus* granted us their amazing assistance and patience. I would also like to thank Brian Binder who was the chief scientist on my first research cruise. His kindness, patience, organization and scientific skills contributed to a wonderful time at sea.

This work would not have existed without the help of Worden lab members, past and present: Adam Monier, Aleja Ortiz, Alyssa Gehman, Darcy McRose, Elif Demir, Fabrice Not, Gus Engman, Harriet Alexander, Heather Wilcox, Heike Möhlig, Jason Hilton, Joshua Lake, Melinda Simmons, Sarah McDonald Dostal, Rajee Ramamurthy, Rory Welsh, Rudolf Gausling and Sebastian Sudek. I cannot thank you enough for your help with the cruises and research, your technical advice, your support, joy and humor. You

have made every day in the lab a little more fun, a little more exciting and a little more inspiring. I will miss you all!

Finally, I would like to express my gratitude to my boyfriend, Mutlu, who encouraged me to embark onto the Ph.D. adventure and supported me throughout the process. My parents, Edith and Antoine, my siblings, Sandrine, Sarah, Magali, Grégory and Marc, as well as the rest of the family who also encouraged me along this long road. All my friends, especially Cédric, Claire, Clément, Clémence, Florence, Jeff, J-O, Joel, John, Joshua, Julie, Kara, Kathy, Kristen, Laurent, Marie, Mark, Martha, Matthieu, Monique, Perrine, Ricardo, Tania, Tammy and Tina have all contributed in small or great ways to this research. Thank you for your moral support, thank you for patience and understanding, and thank you for the fun and relaxing time that I have and will continue to share with you.

Funding for this research was provided by the National Science Foundation, the Gordon and Betty Moore Foundation and the Monterey Bay Aquarium Research Institute. Other support came from the Harding Michel Memorial Fellowship and the Royal Caribbean Fellowship.

TABLE OF CONTENTS

Page

LIST OF FIGURES vi

LIST OF TABLES viii

Chapter

1 INTRODUCTION 1

2 WIDESPREAD DISTRIBUTION OF A UNIQUE MARINE PROTISTAN LINEAGE 13

3 ECOLOGY OF UNCULTURED PRYMNESIOPHYTES 46
Supplementary Information Section I 86
Supplementary Information Section II 99

4 PHOTOACCLIMATION AND ADAPTATION IN THE PICOEUKARYOTIC PRASINOPHYTE *MICROMONAS* 117

5 CONCLUSION 150

LITERATURE CITED 158

LIST OF FIGURES

Page

Chapter 2

Figure 2.1 19

Figure 2.2 24

Chapter 3

Figure 3.1 52

Figure 3.2 75

Figure 3.3 82

Supplementary Information Section I

Figure S3.1 98

Supplementary Information Section II

Figure S3.2 106

Figure S3.3 107

Figure S3.4 108

Figure S3.5 109

Figure S3.6 110

Figure S3.7 111

Figure S3.8 113

Figure S3.9 114

Chapter 4

Figure 4.1 129

Figure 4.2 131

Figure 4.3 133

Figure 4.4 134

Figure 4.5 136

Figure 4.6 137

Figure 4.7 139

Figure 4.8 141

LIST OF TABLES

Page

Chapter 2

Table 2.1 18

Chapter 3

Table 3.1 64

Table 3.2 80

Table 3.3 81

Table 3.4 84

Supplementary Information Section I

Table S3.1 92

Table S3.2 93

Table S3.3 96

Chapter 4

Table 4.1 128

Chapter 1:

Introduction

Marine microbes play an essential role in the global biogeochemical cycles, as they thrive in all aquatic environments. The phytoplankton, (from the Greek terms "phyton" or plant and "planktos" or wanderer) are phototrophic organisms utilizing solar energy to convert CO_2 to organic carbon through photosynthesis. These algae are therefore distributed throughout the euphotic zone waters (i.e. the 'well-lit' surface layer of the ocean), up to 200 m deep. In these surface waters, microbes are heavily involved in the flux of carbon, including the uptake of CO_2 during photosynthesis; the recycling of carbon by respiration; the channeling of fixed carbon to higher trophic levels in the food chain. Furthermore, marine microbes are implicated in the 'biological pump', a process by which carbon is exported from the euphotic zone to the deep ocean.

On a global scale, marine phytoplankton contributes to almost half of the net primary production (Field et al., 1998). Despite this essential role, these organisms form only less than 1% of the Earth total photosynthetic biomass (Simon et al., 2009). This is because marine microbial communities are very dynamic and the biomass is constantly controlled by various factors such as grazing, viruses, nutrient limitations or light. The phytoplankton is composed of a variety free-living, single-celled (on occasion, colonial) organisms ranging from minuscule cyanobacteria to microscopic algae. The number of described species is low, less than 15,000 species of algae (including cyanobacteria and multicellular algae) are described, greatly contrasting with the approximate 300,000 species of terrestrial plants described today (Chapman, 2009, Simon et al., 2009).

However, it has recently become evident that eukaryotic phytoplankton is composed of an array of diverse organisms, many of them not present in culture collections and very small in cell size (Vaulot et al., 2008, Worden and Not, 2008).

The picoplankton

In oligotrophic (nutrient-poor) environments, especially in open oceans, the phytoplankton is often dominated by small organisms, which constitute the base of the marine food chain. Originally, the terms 'picoplankton' (cells <2 μm in size) described minute bacterioplankton (Sieburth et al., 1978). Later, the definition was modified to include all bacteria, archaea as well as eukaryotes in this size fraction (Johnson and Sieburth, 1982). Today, picoplankton often extends (depending on the author) to organisms passing through a 3 μm pore size filter. The smallest picophytoplankton is 0.6 μm in size, while its eukaryotic counterpart measures a mere 0.95 μm in diameter (Raven, 1998, Derelle et al., 2006). Small size seems to be an essential element to the success of picoplankton. The reduced surface to volume ratio of these tiny cells enhances nutrient acquisition and provides a competitive advantage relative to larger organisms (Raven, 1986, Raven 1998).

The picophytoplankton contributes significantly to primary producer biomass, with estimates ranging between 35 to 73% of the total phytoplankton carbon pool (Partensky et al., 1996, DuRand et al., 2001, Li, 1994). In the equatorial Pacific, 60% of the total chlorophyll *a* (Chl *a*) in oligotrophic surface waters and 45% in nitrate-replete waters (at the deep chlorophyll maximum, DCM) fall within the <1 μm size fraction (Chavez 1989, Mackey et al., 2002). In the equatorial Atlantic, similar numbers have been recorded,

with the picophytoplankton contributing to >60% of the Chl *a* biomass (Pérez et al., 2005). Unfortunately, knowledge about the ecology and relative contributions of different taxa of picoplankton is uneven. The picophytoplankton is composed of three main groups of organisms: two genera of cyanobacteria, *Prochlorococcus* (Chisholm et al., 1988, 1992) and *Synechococcus* (Waterbury et al., 1979), and a third group, the picoeukaryotes. In the last 20 years, much research has been devoted to cyanobacteria in part because in open ocean, they are almost always more abundant than picoeukayotes. In temperate and tropical stratified oligotrophic waters, *Prochlorococcus* numerically dominates (Partensky et al., 1999), as shown for example in the Arabian Sea (Campbell et al. 1998, Brown et al., 1999), the tropical Atlantic (Partensky et al. 1996), the equatorial and tropical Pacific (Binder et al., 1996, Blanchot and Rodier, 1996, Landry et al., 1996, Blanchot et al., 2001) or in the Sargasso Sea (Campbell and Vaulot, 1993). *Synechococcus* often has distribution patterns opposite to *Prochlorococcus* (e.g. Binder et al., 1996; Shalapyonok et al, 2001) but can also be numerically significant and contribute to the phytoplankton biomass in tropical and temperate oceans. The picoeukaryotes composing the third main group have been the focus of fewer studies in part due to difficulty in identifying them based on traditional methods.

Picoeukaryotes

In the past, picoeukaryotes were often referred to as small unidentified coccoid cells and flagellates (Booth and Marchant, 1987). This is because the different taxonomic groups in general cannot be accurately separated by flow cytometry (Shalapyonok et al., 2001). In the last decade, it has become evident that small phototrophic eukaryotes, even

though numerically less abundant than marine cyanobacteria, constitute a third active group of the marine picophytoplankton. In the North Atlantic, Li et al. (1992) reported that despite their low abundance, eukaryotic ultraplankton (≈picoeukaryotes) were generally the dominant contributor to carbon biomass in that size fraction. In other oceans of the world, picoeukaryotic importance has also been recognized. In the Arabian Sea, eukaryotes <3 μm in diameter represent 18 to 33% of the total depth integrated phytoplankton biomass (Shalapyonok et al., 2001), while in the equatorial Pacific, they constitute 35% of the picophytoplankton biomass (Mackey et al., 2002). In the Atlantic Ocean, Sargasso Sea and Mediterraneean Sea, they can constitute a substantial part of the picophytoplankton biomass as well (Marañón et al., 2003, DuRand et al., 2001, Brunet et al., 2007). This dominance in terms of biomass is due to their larger cell size relative to other picoplankton (Worden et al., 2004). However, to understand the fate of these organisms, and their role in the carbon cycling, it is critical to move beyond the abundance and standing stock biomass concept and consider primary production as well as carbon transfer to higher trophic levels. Picoeukaryotes can be highly productive in spite of their low abundance; they were reported to be the dominant primary producers within the picoplankton in studies in the North Atlantic, (Li et al., 1992, Li, 1994). In some cases, they can be responsible for up to 76% of the net picoplanktonic production, as shown for a California, USA coastal site (Worden et al., 2004). Furthermore, their role in the food chain is just beginning to be explored. In the same study, picoeukaryotes were subject to higher grazing mortality than *Synechococcus*. This is important because the

transfer of carbon to higher trophic levels allows the material to be packaged into larger particles and potentially contribute to the biological pump whereby organic carbon is removed to the deep ocean.

The importance of picophytoeukaryotes, as well as their phylogenetic diversity is now starting to be appreciated. Exploring diversity is essential in order to create targets for starting to understand the distribution and functions of different groups within a system. For example, the vertical distribution of different genotypes of the cyanobacterium *Prochlorococcus* reflects dramatic variation in the genomic composition of these genotypes and demonstrates a high degree of niche differentiation. Various ecotypes have been shown to be adapted to certain light levels or capable of acclimation to different light levels, reflecting the different light environments encountered in the euphotic zone (Moore, Rocap et al., 1998; Rocap et al., 2003; Johnson, Zinser et al. 2006). If such niche adaptation is observed within this genus, we might expect that various eukaryotic algal taxa will also have specific and distinct roles and physiological processes. To rigorously evaluate the roles and niches of picoeukaryotes in different systems, it is necessary to gain better knowledge of their diversity and the biogeographic distribution of specific taxa.

Picoeukaryotes diversity

Picoeukaryotic diversity is still underexplored. There are various causes for this dearth of knowledge: 1) small eukaryotic algae cannot be easily identified with traditional methods 2) many are not easily cultivated in the laboratory and 3) due to their uneven cellular properties, some taxonomic groups preserve better than others. In the past,

eukaryotes were distinguished by their morphologies and ultrastructures through microscopy (Simon et al., 2000). However, prior isolation and cultivation in the laboratory is required. In addition, many picoeukaryotic taxonomic groups do not possess easily discriminated characteristics visible under light microscopy at the current resolution levels (Andersen et al.1996; Johnson and Sieburth 1982). Subsequently, the use of (scanning and transmission) electron microscopy revealed valuable information on additional taxonomic groups that could not be observed by light microscopy. High performance liquid chromatography (HPLC) is another method to study contributions of various phytoplankton groups to biomass, based on the separation of pigments collected from whole seawater (e.g. Andersen et al., 1996). It generally provides the discrimination of relatively broad taxonomic groups. However, while HLPC gives valuable information on a community, it offers only a limited resolution of the extent of its diversity. Moreover, this method relies on knowledge of pigments from cultured organisms; however, many taxa remained unculturable.

The introduction of molecular techniques to microbial ecology has allowed researchers to gain higher resolution insights to phytoplankton diversity. In particular, the amplification of the plastid-encoded 16S ribosomal RNA (rRNA) from eukaryotes (Rappé et al., 1995, Rappé et al., 1998), then later 18S rRNA gene clone libraries from <2-3 μm size-fractioned environmental samples have unveiled a tremendous and until then unsuspected picoeukaryotic diversity (Moon-van der Staay et al., 2000; Moon-van der Staay et al., 2001; Díez et al., 2001; López-García et al., 2001). The nuclear-encoded 18S rRNA gene is one of the most commonly used markers as it is present in all eukaryotic organisms. In addition, the slow rate of evolution of this gene allows

phylogenetic comparison of distantly related organisms (Vaulot et al., 2008). Finally, the highly conserved regions present in the 18S rDNA are an adequate target for the design of primers suitable for a large amount of eukaryotes. Construction of environmental clone library using 18S rDNA revealed new picoeukaryotic lineages phylogenetically distant to known taxa (e.g. Guillou et al., 1999, Massana et al. 2004b, Not, Valentin et al., 2007). These findings confirmed that picoeukaryotes present in the environment are indeed extremely diverse and that certain lineages still remain uncultured.

Photosynthetic picoeukaryotes belong to many divisions and classes such as Chlorophyta (class Prasinophyceae), Heterokontophyta (class Pelagophyceae, Bolidophyceae, Bacillariophyceae, Chrysophyceae), Haptophyta (class Prymnesiophyceae) (e.g. Vaulot et al., 2008, Worden and Not, 2008). Picophytoeukaryote sequences from various known taxa can be retrieved from environmental samples. For example, at a French coastal site, sequences belonging to the class Prymnesiophyceae, Cryptophyceae, as well as many sequences affiliated to the class Prasinophyceae and, in particular, the order Mamiellales were recovered (Romari and Vaulot, 2004). The diversity of picoeukaryotes has also been studied in other areas, for example in the Arctic (Díez et al., 2001; Lovejoy et al., 2006), in the Mediterranean sea (Massana et al., 2004a, Marie et al., 2006), at a coastal Pacific site (Worden, 2006), in the western Pacific (Cheung et al., 2008), in the Sargasso Sea (Not et al., 2007), in the Arabian Sea (Fuller et al., 2006) and in some cases led to the discovery of unknown taxa or lineages including radiolarians (Not et al., 2007), biliphytes (Not et al., 2007, Not,

Valentin et al., 2007, Cuvelier et al., 2008), a novel class Bolidophyceae (Guillou et al., 1999), or new clades of prasinophytes (Guillou et al., 2004, Worden, 2006, Viprey et al., 2008).

The effect of high light and UV radiation on phytoplankton

The factors influencing picoeukaryotic phytoplankton community composition are still not well understood. Most phytoplankton live in very variable environments and by definition are incapable of large scale directional movement. Cells therefore need to adjust to the conditions of their environments and are often subject to stress imposed by relatively rapid changes (MacIntyre et al., 2000). Among the major factors regulating all photosynthetic plankton growth, light is vital. Yet, excess light and ultra-violet (UV) radiation penetrating the euphotic zone can have deleterious effects on marine life. Because O_3 strongly absorbs UV-B radiation (280-320 nm), depletion of the stratospheric ozone layer has and can result in an increase of the amount of UV-B reaching Earth surface (Meador et al., 2002, Xue et al., 2005). This increase in UV-B has important consequences for marine ecosystems (see for e.g.: Nahon et al., 2008) as UV radiation are known to penetrate the water column to depths of at least 20-30 m in the clearest oligotrophic waters (Kirk, 1994, Tedetti and Sempéré, 2006). In the euphotic zone, high light (HL) also has deleterious effects if energy is absorbed beyond that of the photosynthetic system capacities (Bei-Paraskevopoulou and Kloppstech, 1999). UV radiation and HL can negatively affect phytoplankton growth, survival, pigmentation, metabolism, and photosynthesis (Xue et al., 2005) and therefore primary production, as well as higher trophic levels. If not repaired, damages induced by HL and UV radiation

result in the decrease of photosynthetic capacities of algae, a process called photoinhibition. Under photoinhibition, the electron transport in the photosystem (II) decreases (Tevini et al., 1991). Furthermore, under excessive light, reactive oxygen species are produced in the chloroplast (Hutin et al., 2003). These can cause oxidative damages and irreversible effects to the various components of the photosynthetic apparatus (Norén et al., 2003, Niyogi, 1999). In order to protect themselves against HL and UV radiation, reduce the negative effects of oxidative processes or prevent the formation of oxygen reactive species, photosynthetic organisms have developed various photoprotection or acclimation mechanisms (Niyogi, 1999). These mechanisms include the production of protective pigments and proteins, and the alteration the photosystem apparatus composition (Salem and van Waasbergen, 2004). For example, the absorption and utilization of light energy in the chloroplasts can be regulated. In addition, the amount of photosynthetic pigments and photosynthetic efficiency can be adjusted (Niyogi, 1999).

The photosystem and photoprotection genes

During photosynthesis, solar energy is transferred to the photosystem (PS), which is composed of antennae (also called light harvesting complexes, LHC) and a reaction center (RC). The antennae which capture light and transfer the energy to the RC consist of one or more chromophores (small molecules that absorb light) bound to one or more proteins. In eukaryotes, chromophores can be chlorophylls, phycobilins and carotenoids and the proteins to which they are usually bound can belong to seven major proteins families, one being the LHC protein family (Green et al., 2003). *Lhc* genes are nuclear

encoded genes and are translated in the cytosol. Their products are then transported to the chloroplasts where they bind to the pigment and are inserted in the thylakoid membranes of the chloroplasts (Koziol et al., 2007). In land plants and green algae, LHC proteins belong to the Chl *a/b*-binding proteins, while in red algae and chromalveolates, they belong to the Chl *a*-binding proteins and Chl *a/c*-binding proteins, respectively (Koziol et al., 2007). Most photosynthetic organisms possess two photosystems: PSI and PSII with LHCI and II as their respective LHC proteins; while the structure and function of PSII is well understood, PSI is less known.

In prasinophyte algae, the structure of PSII is very similar to plants with a few exceptions. However, there are several noticeable differences in pigment-protein complexes in *Mantoniella squamata*, the first representative of the class Prasinophyceae (order Mamiellales) to be explored. Studies have shown that only one unique LHC type, named LHCP ("P" for prasinophyte), was present in this organism. This is interesting because it is hypothesized that in ancestral algae a single LHC was associated with both photosystems; hence *Mantoniella* might represent such an ancestral state. However, the presence of LHCI genes and proteins was revealed by Six, Worden and colleagues (2005) in *Ostreococcus tauri* (order Mamiellales, class Prasinophyceae) in addition to the LHCP. In this work the authors returned to *M. squamata* and showed that the earlier work had simply "missed" LHCI proteins in that organism due to the use of overly specific antibodies. Therefore, even though the LHCP proteins were abundant LHC proteins in Mamiellales, LHCI proteins also formed a significant fraction of the LHC proteins. These findings raise questions about the specific function of LHCP and potential specific adaptations of prasinophytes to a range of light fields, intensities or variations. Even

though it is likely that LHCP are associated with PSII, the phylogenetic position of LHCP proteins outside the clade containing the LHCII polypeptides of plants and green alga, make the Mamiellales a unique model (Koziol et al., 2007). The picoeukaryotic green algae *Micromonas sp.* strain RCC299 (Mamiellales) possesses four *Lhcp* genes (*Lhcp1-4)*, one of which (*Lhcp2*) has seven copies (Worden et al., 2009). In this dissertation, the expression of *Lhcp1* transcripts under HL and UV light is investigated. In addition, *Micromonas* possess genes encoding proteins with similarities to the LHC protein family. These proteins are referred as the light-harvesting-like (LIL) proteins (Jansson, 1999) but their nomenclature is not well established and their names differ depending on the organisms in which they are present or the conditions under which they were first described. LIL proteins possessing a) one transmembrane helix are often referred as one-helix proteins (OHPs), also called high light induced proteins (HLIPs) in cyanobacteria (Dolganov et al., 1995), b) two helices are often named stress-enhanced proteins (SEPs) in plants (Heddad and Adamska, 2000) and c) three helices are often designated early light induced proteins (ELIPs, Adamska, 1997) but also included LHCSR (stress-related members of the LHC protein family (Peers et al., 2009), formerly called LI818, (Gagné and Guertin, 1992). All these proteins are thought to have a role in stress responses, including excess light stress. However, their specific roles remain unclear. Here we will investigate the expression of two of those genes often referred to as *Ohp2* and *Lhcsr*, although other names are used as well.

The objectives for chapter 2 of the dissertation were to:

1) Explore the molecular phylogenetic diversity of a novel group of picoeukaryotes, namely the biliphytes, in tropical (Florida Straits) and subtropical (Sargasso Sea) environments.
2) Estimate biliphytes characteristics and abundance.
3) Determine their potential contribution to the phytoplankton biomass, based on the assumption that they are photosynthetic.

The objectives for chapter 3 of the dissertation were to:

1) Investigate the phylogenetic diversity of uncultured pico-prymnesiophytes.
2) Determine the geographic distribution, abundance and biomass of picoprymnesiophytes in the five major ocean basins.
3) Estimate their growth and grazing mortality rates in the Sargasso Sea.

The objectives for chapter 4 of the dissertation were to:

1) Determine the effect of HL and UV radiation on the growth and physiology of the phototrophic picoeukaryote *Micromonas.*
2) Quantify the expression of specific genes (*Lhcp1*, *Ohp2* and *Lhcsr*), hypothesized to be involved in photosynthesis and photoprotection of *Micromonas*, under experimental conditions that would be stressful to these cells, but are environmentally relevant.

Chapter 2:

Widespread distribution of a unique marine protistan lineage

Summary

Unicellular eukaryotes (protists) are key components of marine food webs, yet knowledge of their diversity, distributions and respective ecologies is limited. We investigated uncultured protists using 18S rRNA gene sequencing, phylogenetic analyses, specific fluorescence *in situ* hybridization (FISH) probes and other methods. Because few studies have been conducted in warm water systems, we focused on two Atlantic subtropical regions, the Sargasso Sea and the Florida Current. Cold temperate waters also were sampled. Gene sequences comprising a unique eukaryotic lineage, herein termed "biliphytes," were identified in most samples, whether from high (30 ^{o}C) or low (5 ^{o}C) temperature waters. Sequences within this uncultured group have previously been retrieved from high latitudes. Phylogenetic analyses suggest biliphytes are a sister group to the cryptophytes and katablepharids, although the relationship is not statistically supported. Unlike results from the initial publication on these organisms (therein "picobiliphytes"), we could not detect a nucleomorph, either visually, or by targeted primers. Bootstrap supported sub-clades were delineated but coherence was not obvious with respect to geography or physico-chemical parameters. Phycobilin-like fluorescence associated with biliphyte-specific FISH probed cells supports the hypothesis that they are photosynthetic. Furthermore, our data indicate the biliphytes are nanoplanktonic in size, averaging from 4.1 ± 1.0 x 3.5 ± 0.8 μm (± SD) for one probed group, and 3.5 ± 0.9 x 3.0 ± 0.9 μm (± SD) for another. We estimate biliphytes contributed 28 (± 1) % of the

phytoplanktonic biomass in tropical eddy-influenced surface waters. Given their broad thermal and geographic distribution, understanding the role these protists play in biogeochemical cycling within different habitats is essential.

Background

Marine protists are vital components in the global carbon cycle. Knowledge of their diversity and respective ecologies is nascent yet rapidly growing as culture independent approaches are applied in a variety of natural settings. Application of polymerase chain reaction (PCR) based approaches in marine systems has revealed a tremendous degree of eukaryotic diversity, using markers such as the plastid encoded 16S ribosomal RNA (rRNA) gene (Rappé et al., 1998) and the RuBisCo large subunit (rbcL) gene (Paul et al., 2000) as well as the nuclear encoded 18S rRNA gene (Díez et al., 2001). Phylogenetic analyses of environmental sequence data have demonstrated the existence of many novel clades within the eukaryotic tree of life, with many sequences from "pico"-size (generally defined as <2-3 μm) fractionated samples (López-García et al., 2001; Moon-van der Staay et al., 2001; Massana et al., 2004a; Groisillier et al., 2006; Not et al., 2007). The majority of these studies have concentrated in 'local' coastal zones (Massana et al., 2004b; Romari and Vaulot, 2004; Countway et al., 2005; Worden, 2006), or extreme environments, such as deep or polar waters (López-García et al., 2001; Lovejoy et al., 2007). Consequently, there is a dearth of knowledge with respect to the diversity and abundance of subtropical and tropical microbial eukaryotes. Furthermore, little is known

about the morphological or functional attributes of the organisms from which such sequences are derived. This is largely due to the fact that there are no cultured representatives for many of the newly identified clades.

Linked investigations of the diversity, abundance and function of uncultured protists are critical to understanding evolutionary and ecological aspects of these populations, as well as for efforts to improve biogeochemical models. For prokaryotes, tremendous advances have been made via ecological genomics, whereby, in some cases it has been possible to begin to assign functional roles to uncultured bacteria (Moran and Miller, 2007). However, to date, we lack insights on uncultured eukaryotes akin to those gained from prokaryotic metagenomic studies (DeLong et al., 2006; Worden et al., 2006; Rusch et al., 2007) and the genomic potential of protistan communities remains virtually unknown. While metagenomic studies provide a first glimpse at possible metabolic capabilities, linking them with rigorous environmental characterization and activity measurements allows basic understanding of ecosystem level processes to advance (Azam and Worden, 2004; Moran and Miller, 2007). For uncultured protists, such approaches promise a means to elucidate their physiologies and ecologies. At this stage identification of novel eukaryotic groups and data their distribution and abundance is needed to facilitate selection of appropriate sampling sites and populations for targeted metagenomics.

Two studies have recently highlighted the presence of a protistan group in high latitude marine surface waters with unknown affinities to other eukaryotes and termed “picobiliphytes” (Not et al., 2007; Not, Valentin et al., 2007). 18S rRNA gene sequences were recovered from the Arctic Ocean, the Norwegian Sea, and irregularly in fall and

winter coastal European waters, suggesting that these organisms are widespread in cold and polar waters (Hearn, 2007; Not, Valentin et al., 2007). However, enumeration of these putatively photosynthetic cells showed they were extremely rare, composing less than 1% of the eukaryotic community (Not, Valentin et al., 2007).

Few studies have explored molecular phylogenetic diversity in subtropical and tropical environments or linked data on novel eukaryotic sequences with organism characteristics or abundance. We undertook a comprehensive set of research expeditions to explore subtropical marine protist communities. Here, we investigate the phylogeny of a unique set of sequences discovered in samples from subtropical regions as well as from cold waters. We show that this enigmatic lineage, here termed "biliphytes", is composed of nanoplanktonic (2-20 μm) organisms falling within genetically distinct clades, that also harbor the previously published "picobiliphyte" sequences (Not, Valentin et al., 2007). Based on results to date, we infer a functional role for the biliphytes and assess their ecological significance in the subtropics.

Material and Methods

Field sites and sample collection

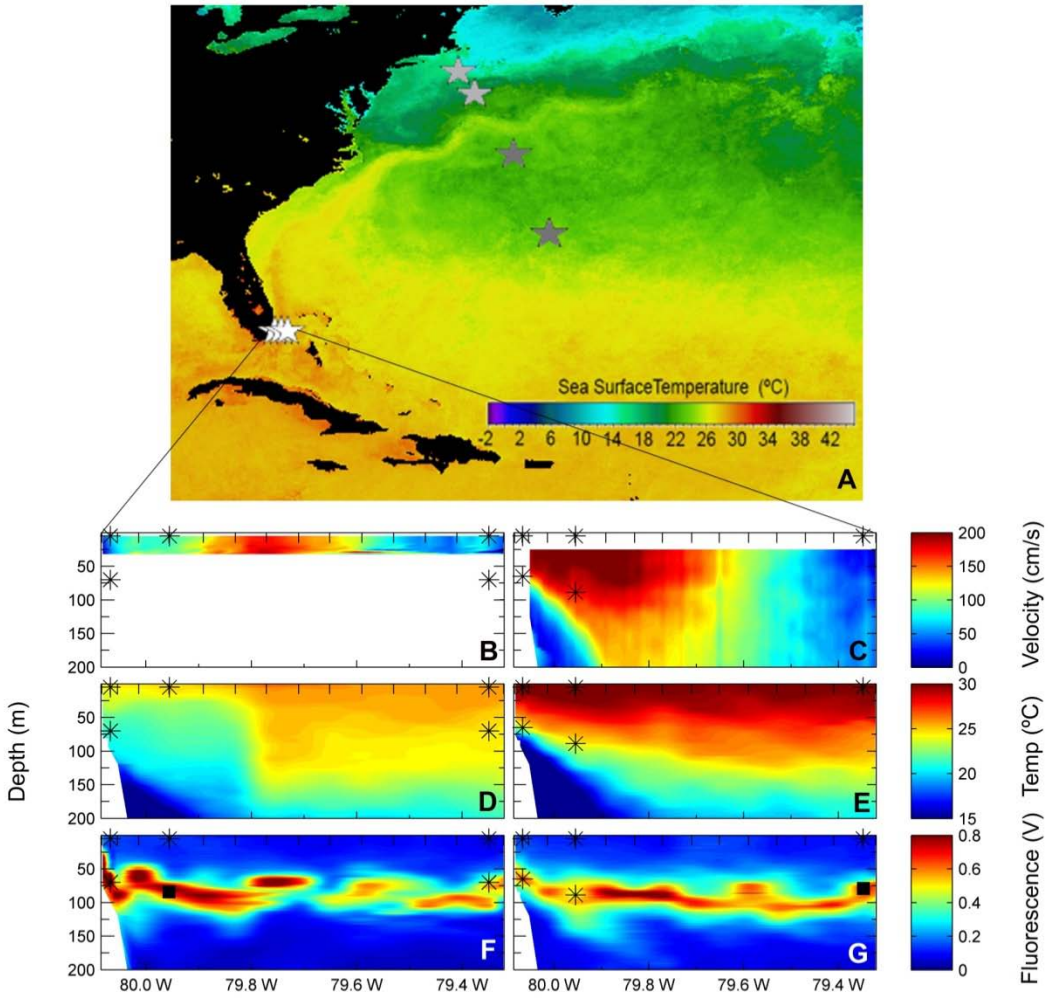
Five cruises were conducted, two transects from coastal New England to the Bermuda Atlantic Time-series Study (BATS) site (R/V Endeavor cruise EN351 and R/V Oceanus cruise OC413) and 3 transects across the Florida Straits (WS0503, WS0518, WS0528) with characteristics as in Table 2.1.

Table 2.1. Coordinates, dates, environmental characteristics and associated 18S rDNA sequences for sample sites. BATS, Bermuda Atlantic Time-series Study Station; NSS, Northern Sargasso Sea; CSS, Continental Shelf-edge/Slope; CS, Continental Slope; FS, Florida Straits, Station 01, South Florida, USA side; Station 04, care Gulf Stream forming waters; Station 14, Bahamas side.

*denotes parameters measured with CTD detector prior to the sample collection (if a single measurement) or prior to and the following day (if a range of measurements), as samples were collected with a single GO-FLO bottle not equipped with a CTD detector. N/A: data not available.

Site	Latitude (N)	Longitude (W)	Date (d/m/y)	Sample Depth (m)	Bottom Depth (m)	Temperature (ºC)	Salinity (ppt)	Sequence names
BATS	31º 39' 20"	64º 37' 21"	05/29/05	75	4397	20.08*	36.79*	OC413BATS_O043_75m
BATS	31º 39' 20"	64º 37' 21"	06/01/05	15	4397	25.51*	36.68*	OC413BATS_P006_15m OC413BATS_P082_15m
NSS	35º 09' 24"	66º 33' 46"	06/05/05	15	4980	21.37-21.56*		OC413NSS_Q007_15m OC413NSS_Q040_15m OC413NSS_Q086_15m
CSS	40º 15' 07"	70º 25' 23"	04/09/01	20	706	5.31	32.71	EN351_CTD040_20m
CSS	40º 15' 07"	70º 25' 23"	04/09/01	4	706	5.60	32.70	EN351 CTD040 4m
CS	39º 10' 55"	69º 31' 08"	04/10/01	30	2750	14.24	35.78	EN351_CTD039_30m
FS01	25º 30' 07"	80º 04' 04"	03/30/05	5	125	24.26	36.43	FS01B026_30Mar05_5m FS01B048_30Mar05_5m FS01B033_30Mar05_5m FS01B029_30Mar05_5m
FS01	25º 30' 07"	80º 04' 04"	03/30/05	70	125	21.15	36.41	FS01C040_30Mar05_70m
FS01	25º 30' 04"	80º 03' 59"	08/01/05	5	100	30.08	36.10	FS01AA11_01Aug05_5m FS01AA94_01Aug05_5m
FS01	25º 30' 04"	80º 03' 59"	08/01/05	65	100	23.51	36.32	FS01D014_01Aug05_65m FS01D065_01Aug05_65m FS01D057_01Aug05_65m FS01D054_01Aug05_65m FS01D031_01Aug05_65m FS01D022_01Aug05_65m
FS04	25º 30' 01"	79º 57' 20"	03/31/05	5	375	24.68	36.31	FS04E037_31Mar05_5m FS04E081_31Mar05_5m
FS04	25º 30' 04"	79º 57' 18"	08/01/05	5	350	30.29	36.01	FS04GA46_01Aug05_5m FS04G190_01Aug05_5m FS04G188_01Aug05_5m FS04GA95_01Aug05_5m
FS04	25º 30' 04"	79º 57' 18"	08/01/05	89	350	24.22	36.54	FS04H169_01Aug05_89m FS04H153_01Aug05_89m
FS14	25º 29' 59"	79º 20' 58"	03/30/05	5	729	25.65	36.20	FS14JA65_30Mar05_5m FS14JA72_30Mar05_5m
FS14	25º 29' 59"	79º 20' 58"	03/30/05	70	729	24.65	36.61	FS14I06_30Mar05_70m
FS14	25º 29' 59"	79º 20' 54"	07/31/05	5	650	29.88	36.04	FS14K017_31Jul05_5m FS14K025_31Jul05_5m
FS14	25º 30' 01"	79º 21' 04"	12/08/05	58	N/A	26.18	36.26	FS14M008_08Dec05_58m FS14M021_08Dec05_58m

Figure 2.1. Geographic location, physical and biological parameters of the sample regions and sites. Three regions were investigated during the cruises. In the Sargasso Sea (dark grey stars), 2 sites were sampled: BATS (bottom star) and the Northern Sargasso (top star). In the continental shelf area (light grey stars), 2 sites were sampled: the continental edge (bottom star) and the continental slope (top star). In the Florida Straits (white stars), 3 sites were sampled for each of the three cruises (March 05, July/August 05, December 05): Station 01 (left star), Station 04 (middle star) and Station 14 (right star) (A). Bottom panels show the cross-straits vertical profiles of the north component of current during March (B) and July (C); temperature during March (D) and July (E); and fluorescence during March (E) and July (F). Sampling stations and depths containing biliphytes are indicated with an asterisk. Sampling stations and depth where biliphytes sequences were not recovered are not shown. Tick marks on the upper x-axis indicate stations at which CTD casts were performed to measure environmental parameters. During the March cruise the ship was only equipped with a 600 kHz ADCP capable of measuring shallow currents, hence (B) only shows velocity in the upper 30 m of the water column. SeaWIFS data (A) is derived from http://oceancolor.gsfc.nasa.gov/cgi/browse.pl from the integrated June 2005 sea surface temperature data.



Samples were collected in surface and deep chlorophyll maximum (DCM) waters using a Sea-Bird Niskin Rosette equipped with standard CTD and PAR detectors (see also below). For DNA extraction samples, 1 l of seawater was gravity filtered through 2 µm pore size filters (GE Osmonics, Minnetonka, Minnesota, USA) and then onto a 0.45 µm pore size (2001) or

0.2 µm pore size (2005) Supor filter (Pall Gelman, Ann Arbor, Michigan, USA) using vacuum. For FISH samples whole seawater (no size fractionation) was preserved with 1% paraformaldahyde (final concentration) for at least 1 hour in the dark at 4 °C. For each replicate, volumes of 180 ml or more of seawater were gently filtered onto a 25 mm, 0.2 µm Anodisc filter (Whatman, Maidstone, England). The filters were then subjected to an ethanol dehydration series at 50%, 80%, 100% (diluted in sterile, 18.2 ΩO H_2O) for 3 min each and stored at -80 °C. Flow cytometry (FCM) samples were preserved with 0.25% (final concentration) fresh electron microscopy grade glutaraldehyde (Tousimis, Rockville, Maryland, USA), flash frozen in liquid nitrogen and moved to -80^{o} C for long term storage.

DNA extraction and PCR amplification

DNA was extracted using the DNeasy kit (Qiagen, Germantown, MD, USA). 18S rRNA genes were amplified using primers (5′- ACCTGGTTGATCCTGCCAG-3′ and 5′-TGATCCTTCYGCAGGTTCAC-3′) complimentary to conserved regions proximal to the gene termini, designed as universal eukaryotic primer set, but likely with some biases (Moon-van der Staay et al., 2001). PCR was performed with an initial “hot start” for 15 min at 95 °C, proceeded by 32 cycles at 94 °C for 30 sec, 55 °C for 30 sec, and 72 °C for

1 min; followed by a final extension at 72 °C for 15 min. One µl of PCR product was ligated into the vector pCR2.1 (Invitrogen, Carlsbad, California, USA) and transformed, after colony selection and growth plasmids were purified. Sequencing protocols were based on the di-deoxy sequencing method. Multi-well cycle-sequencing reaction plates were prepared with plasmid template DNA and sequencing reactions were completed using the Big Dye Terminator chemistry and sequenced with either two plasmid targeted primers (M13F and M13R), with the former two primers and primers internal to the PCR product, 1174R and 502F, or with just the latter (Worden, 2006). Reaction mixtures, thermal cycling profiles, and electrophoresis conditions were optimized to reduce the volume of the Big Dye Terminator mix (Applied Biosystems, Foster City, CA, USA) and to extend read lengths on the AB3730xl sequencers (Applied Biosystems, Foster City, CA, USA). 96 clones were sequenced for each of the 17 libraries (14 of which yielded biliphyte sequences) constructed from 2005 cruise samples. For 2001 samples 48 clones were sequenced for each of the 3 libraries, all of which yielded biliphyte sequences.

Primers were also designed to preferentially amplify red algal and/or cryptophyte nucleomorph 18S rRNA sequences (NMR-SSU-1102-F: 5′-GAAATCAAAGTSTHTGGGTTCT-3′; NMR-SSU-1607-R: 5′-TACAAAGGGCAGGGACGTAT-3′; NM-SSU-0465-F: 5′-AATCCTGAYTMAGGGAGGTAGC-3′; NMR-SSU-0881-R: 5′-TCACCTCTGRCAAYGRARTAC-3′; NMR-SSU-0150-F: 5′-CTACKTGGATAMCCGTAG-3′; NMR-SSU-0748-R: 5′-GCYWTGAACACTCTAHTTTNTTC-3′). PCR and nested PCR was performed with these primers using 30 cycles of 95 °C for 30 sec/1 min, 45 °C for 1 min and 72 °C for 30

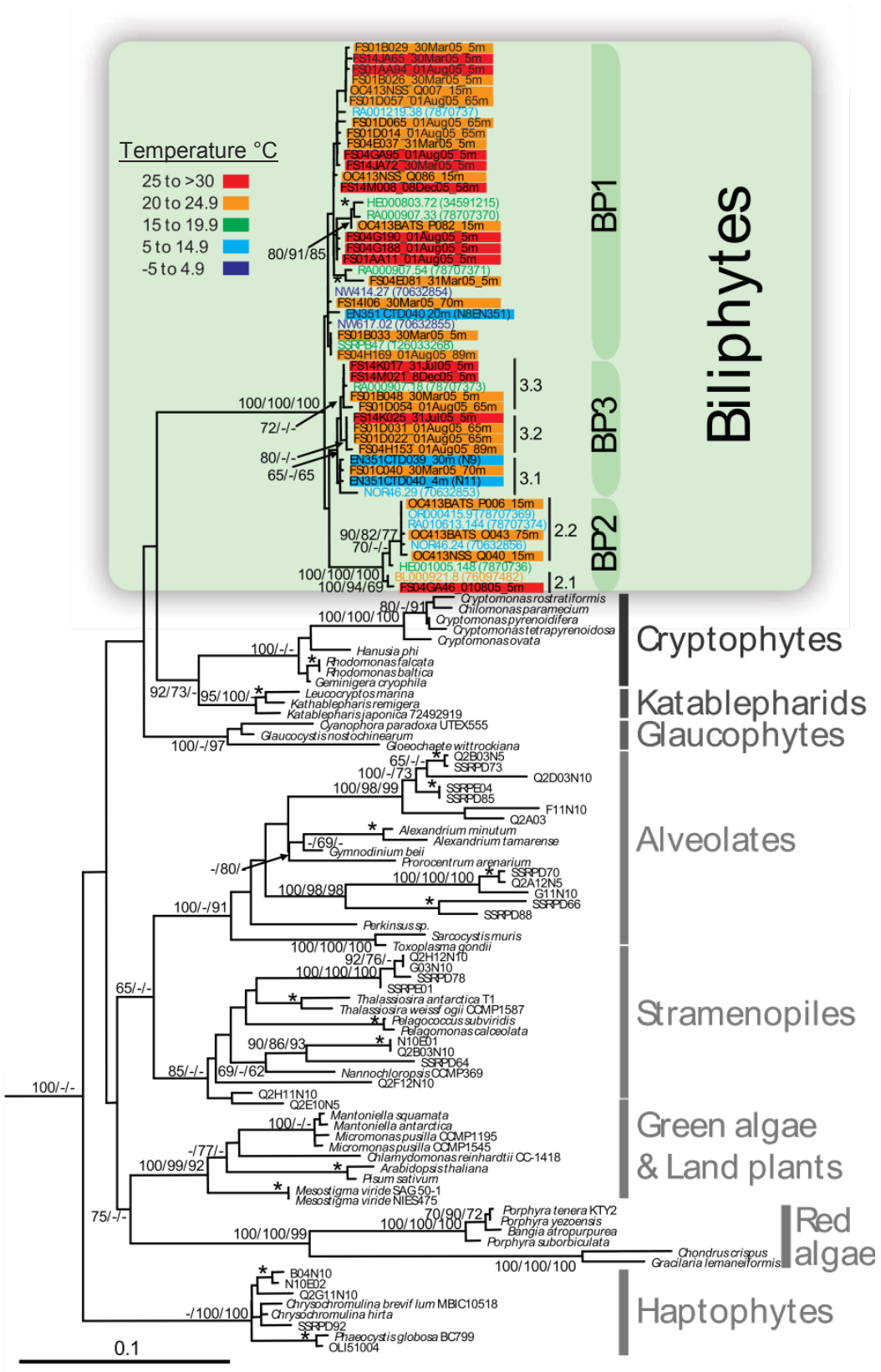
sec/1 min. PCR products were cloned using the TOPO-TA cloning kit (Invitrogen, Carlsbad, CA, USA) as per manufacturer's instruction and sequenced on a Beckman Coulter CEQ8000 (Beckman Coulter Inc., Fullerton, California, USA).

Phylogenetic analyses

BLASTN and preliminary phylogenetic analysis was used to identify biliphyte-like sequences in the Genbank non-redundant (May 2006, March 2007) and CAMERA (April 2007) databases. It became apparent that several deposited clones (Not, Valentin et al., 2007) had been deposited previously (Romari and Vaulot, 2004) under other accession numbers. Here only one sequence version per clone is used. CHECK_CHIMERA (Maidak et al., 2001) and manual screening were used to identify and remove likely chimeras as necessary. The alignment used for final phylogenetic analyses was performed in ClustalW. The alignment was then visual curated, ambiguously aligned sites were removed or re-adjusted (manual alignment) in obvious cases where ClustalW failed. Finally, the alignment was masked to use only those bases within regions of unambiguous alignment and considered homologous. Prior to this, Chimera Check (RDP II) and manual screening of alignments were used to identify and remove likely chimeras. Several short originally unidentified sequences from a previous publication (Countway et al., 2005) were not used in the alignment, none of which fell in BP2. Model selection, number of rate categories, proportion of invariable sites, transition/transversion (TiTv) and the gamma distribution parameter were determined within Modeltest (Posada and Crandall, 1998). Phylogenetic analysis was performed using maximum likelihood, distance neighbor joining methods and parsimony, all within the Phylip modules

(Felsenstein, 2005). For maximum likelihood 100 datasets were used for bootstrapping, with global rearrangements, randomized input, 10 jumbles and 6 categories (fraction of invariable sites = 0.2332; TiTv = 1.9433; $1/\alpha = 1.451258$). The same alignment and parameters were used for neighbor-joining distance (1000 replicates, 10 jumbles) and parsimony analysis (100 replicates, 2 jumbles).

Figure 2.2. Phylogenetic analysis of 18S rRNA gene sequences from environmental clone libraries (all non-italic sequence names) and cultured representatives of eukaryotic first rank taxa. Sequences in colored bars are reported for the first time in this study and are derived from 17 environmental clone libraries, 3 from the Sargasso Sea (prefixes: OC413BATS_P, Q; OC413NSS), 11 from 3 stations in the Florida Straits (prefixes: FS01, FS04, FS14) over multiple dates and depths (see Table 2.1), as well as 3 from continental shelf-edge and slope waters (prefixes: EN351). Biliphyte sequences from previous work are shown in colored font. Bar and font colors correspond to 5 temperature ranges as indicated on the figure. Other Sargasso Sea and Pacific Ocean environmental sequences (non-italic, black font sequence names without colored bars) are included to illustrate the distinct nature of biliphyte sequences relative to other uncultured eukaryotes retrieved from the same waters. The tree shown was inferred by maximum likelihood (ML) methods using the model TrN + I + Γ (I = 0.2332; Γ = 0.4748) with global rearrangements, randomized sequence input, 10 jumbles, Transition/Transversion = 1.9433 and 6 rate categories. Node values reflect bootstrap support in the order ML/Neighbor-joining Distance/Parsimony as percentages of 100/1000/100 replicates. Bootstrap support for terminal nodes (*) is indicated only if over 90% by all 3 methods. Two radiolarian sequences (126033309 and 126033226) served as an outgroup (not shown). The scale bar indicates the estimated number of nucleotide substitutions per site.


Temperature °C
25 to >30
20 to 24.9
15 to 19.9
5 to 14.9
-5 to 4.9
Biliphytes
BP1
BP3
BP2
3.3
3.2
3.1
2.2
2.1
Cryptophytes
Katablepharids
Glaucophytes
Alveolates
Stramenopiles
Green algae & Land plants
Red algae
Haptophytes
0.1

Cell enumeration, measurements and biomass determinations

Whole cell 18S rRNA targeted FISH was performed on replicate filter pieces in conjunction with tyramide signal amplification (TSA) using a modification of a previously published method (Not et al., 2002) and newly designed biliphyte-targeted probes targeting BP1 (5´-GCGTGATGCCAAAATCCG-3´) and BP2 (5´-ATATGCCCGTCAAACCGT-3´), the two probes originally verified and used in Not, Valentin et al. (2007). For the hybridizations, the filters were partitioned into multiple pieces for use with different probes. One µl of probe stock solution (50 ng probe μl^{-1} stock solution) and 9 µl of 40% hybridization buffer were mixed and added to each filter piece, then incubated for 3 h in a humid chamber. In order to make the hybridization buffer, a solution of blocking reagent (150 mM NaCl, 100 mM maleic acid, 10% blocking reagent, w:v) was first prepared. The final concentrations for the hybridization buffer were: 40% deionized formamide, 20 mM Tris-HCl, pH 7.5, 0.01% sodium dodecyl sulfate, 2% blocking reagent (w:v), 20 mM maleic acid, 930 mM NaCl (final concentration including the NaCl used for the blocking reagent). Following the incubation, the samples were washed twice for 20 min at 37 °C (5 mM EDTA, 0.01% SDS, 20 mM Tris, 37 mM NaCl) and equilibrated in TNT buffer (100 mM Tris HCl, pH 7.5, 150 mM NaCl, Tween 20) at room temperature (RT) for 15 min. The TSA Fluorescein Tyramide Reagent Pack (PerkinElmer, Boston, MA, USA) was used for amplification of the signal. Ten µl of kit reagent mix (1:1 amplification diluent and 40% dextran sulfate, 1:50 Fluorescein Tyramide) was added to the filter and incubated at RT in the dark for 30 min. The reaction was stopped by 2 consecutive baths (20 min each) in TNT buffer at 55 °C. The filter pieces were equilibrated in nanopure water at RT for 7

min in the dark and counterstained with 4', 6-diamidino-2-phenylindole (DAPI, 2.5 µg ml^{-1}) for 5 min. Finally, the filters were rinsed 5 min at RT in nanopure water and briefly dipped in 80% ethanol. Seven µl of mounting solution (1:5 antifading solution AF1, Citifluor, London, UK and VECTASHIELD mounting medium, H-100, Vector Laboratories, Burlingame, CA, USA) was added to each filter piece and a coverslip was fixed with nail polish. The filters were counted within 24 hours. Probe specificity was initially verified using cultures of non-target species: *Rhodomonas salina* (CCMP1319), *Isochrysis* cf sp. (CCMP1244), *Micromonas pusilla* (NOUM17), *Ostreococcus tauri* (OTH95), and *Paraphysomonas imperforata* to ensure that these cells were not illuminated by the biliphyte probes. The bacterial anti-sense NON338 probe was used as a negative control for all hybridizations. In addition to procedures used in Not, Valentin et al. (2007), two negative controls (BP1 and BP2 probes on cultured *R. salina* CCMP1319 cells; and NON338 on all field samples) were employed alongside hybridization of all field samples as well as a no-probe control. 50 fields were enumerated per sample using a 100X oil-immersion objective on an Olympus BX61 epifluorescence microscope, probe signal in the FITC channel and associated DAPI fluorescence was enumerated. For cell sizing a total of 105 cells were measured using a calibrated sizing grid and reported as the average shortest dimension by the average longest dimension.

Florida Straits physical and fluorescence characteristics analyses

At each of the 16 Florida Straits station physical measurements were made through the entire depth of the water column using a CTD equipped with Seabird sensors measuring

temperature, salinity, fluorescence, oxygen, beam transmission and light levels. Continuous measurements of the currents were made using a 600 kHz (samples 2 m depth bins from 4-35 m) Acoustic Doppler Current Profiler (ADCP; Teledyne RD Instruments, Poway, CA, USA) in March and a 75 kHz (samples 16 m depth bins from 25-400 m) ADCP in July. All physical data were processed using the manufacturer provided software, and were plotted using Matlab (Mathworks Inc., Natick, MA, USA), with a cubic interpolation between stations. Some temporal aliasing is introduced into the plots due to the sampling of the eastern half of the transect (e.g. Stations 07-15 in March 2005) on the first night of the cruise, and the western half (e.g. Stations 06-00) of the transect on the second night of the cruise. In July, sampling of the eastern half of the transect (e.g. Stations 08, 10-15) occurred on the first night, and other stations (e.g. Stations 09, 07-00) on the second night, of the cruise.

Results

Study sites

To attain broadly representative results we investigated molecular diversity in an array of marine samples, and population characteristics in a subset of those. Three regions were explored during cruises undertaken in 2001 and 2005, two regions being subtropical and one a cold water environment (Figure 2.1, Table 2.1). The first region, the Sargasso Sea, is an open-ocean subtropical Gyre system. Two sites were sampled within this area, BATS station and a location in the northern Sargasso Sea. At the time of sampling the water column at BATS was under summer-like conditions, with the mixed layer extending to 30 m. At the Northern Sargasso station, the mixed layer extended slightly

deeper, to 40 m. Water samples were collected from the DCM, located well below the seasonal thermocline, and from close to the surface (Table 2.1). The second region, the Florida Straits, is located between south Florida, USA and the Bahamas. The Florida Current, which separates the Florida Keys, USA from Cuba and connects the Gulf of Mexico with the Atlantic Ocean, flows at high velocity (Figure 2.1B and 2.1C) through the Straits, and composes a large fraction of the Gulf Stream forming waters. At the latitude sampled, this Current is in many aspects representative of tropical waters: it is close to the tropics, which spans the equator from 23° 27' N to 23° 27' S, in latitude (our transect being at 25 °C 30') and maintains high water temperatures year round, with extended periods around 30 °C, unlike subtropical systems such as BATS which undergo a greater temperature variations. The physical properties of the Straits, atmospheric conditions and current velocity, contribute to retention of tropical water characteristics in particular zones of the Straits (see Figure 2.1 and satellite data at http://imars.usf.edu). This location was sampled intensively, on a series of 3 cruises representing different seasonal and event-driven variations within the system including, the remnants of a cyclonic mesoscale eddy on the western side during spring (Figure 2.1B, 2.1D and 2.1F), the highly stratified summer waters (Figure 2.1C, 2.1E and 2.1G), and the mixing winter waters. Physical and fluorescence parameters were measured at 16 stations distributed from the western side (south Florida) to the eastern side (Bahamas) in order to comprehensively map characteristics of the water mass. Water samples were taken at the three stations which best represent different water zones in the Straits: the western side (Station 01), the more central Gulf Stream forming waters (Station 04) and the eastern side which has a deeper bottom depth (Station 14, Table 2.1). Samples were collected at

the surface and DCM for the three stations on a March/April cruise, and a July/August cruise. Only one sample, from the DCM at Station 14, was utilized from a December cruise due to limited sampling resulting from rough seas. Overall, five of the Florida Straits samples were from high temperature waters, ranging to 30 °C (28.4 ± 2.2 °C, average and SD, Table 2.1). The third region investigated was continental slope waters south east of Cape Cod, USA. Two sites were sampled at the surface and DCM: over the continental slope (bottom depth of 2750 m) and over the continental slope close to the continental shelf boundary (bottom depth of 706 m). Temperatures for these waters samples ranged from 5.3 to 14.2 °C.

Gene Sequences and Phylogenetic Analyses

The unique 18S rRNA gene sequences were retrieved from seventeen out of twenty clone libraries (Figure 2.1, Table 2.1). Phylogenetic analyses (Figure 2.2) of the newly obtained sequences were performed in the context of publicly available data covering the known breadth of eukaryotic diversity. The results delineate a eukaryotic group at the first rank taxon level (Adl et al., 2005). Using an alignment of 126 taxa, a strongly supported clade was identified, containing 18S rRNA gene sequences derived from the subtropical regions, including remnants of a tropical eddy, and the continental slope sites. The "picobiliphyte" sequences previously obtained from cold surface waters of the Arctic Ocean, the Norwegian Sea, as well as fall and winter coastal European waters (Not, Valentin et al., 2007) were scattered amongst these sequences (Figure 2.2). Within this clade, no strong correlation between phylogenetic relatedness and geographic location was apparent, though several highly supported clusters of sequences are apparent (e.g.,

BP2; Figure 2.2). As observed previously (Not, Valentin et al., 2007), our phylogenies resolve the biliphytes as a potential sister group to the cryptophytes and katablepharids. This relationship is, however, not statistically supported, similar to results obtained for most of the backbone of the 18S rRNA gene tree.

In sum, we isolated biliphyte 18S rRNA gene sequences from all five subtropical sites, BATS, the Northern Sargasso Sea and the three Florida Straits stations. Furthermore, populations were detected during all time periods investigated, including waters up to 30 °C and in slope waters as cold as 5.3 °C (Figure 2.1 and 2.2). Biliphyte sequences were retrieved from the surface and the DCM, close to the base of the euphotic zone. It should be noted that three other sequenced clone libraries, generated from samples taken on the same research expeditions, did not contain biliphytes amongst sequenced clones (data not shown). These samples were from the Florida Straits DCM at Station 04 in March and Station 14 in July as well as the DCM in the Northern Sargasso Sea Station. No biliphyte-like 18S rRNA gene sequences were identified within the GOS data or CAMERA database.

Characteristics and abundance

Because cellular abundance cannot be inferred based on clone library numbers and function cannot be assigned based on 18S rRNA gene phylogeny, water samples were interrogated using two biliphyte-specific, rRNA-targeted FISH probes. Probes targeted a subset of sequences that fell within clades BP1 and BP2 (Figure 2.2). These revealed a phycobilin-like autofluorescence (orange) within the fluorescein-labeled, probed cells, indicating that these organisms may be photosynthetic in nature, similar to the findings of

Not, Valentin et al. (2007). Chlorophyll-like autofluorescence was not observed, nor expected, given that it is an alcohol soluble pigment and filters were treated with an alcohol dehydration series prior to FISH.

Cells identified using the specific FISH probes fell within the nanoplankton size class (2-20 μm). Those hybridized by the BP1 probe were slightly larger than those with the BP2 probe. The former averaged 4.1 ± 1.0 x 3.5 ± 0.8 μm (± SD, n = 60), while the latter averaged 3.5 ± 0.9 x 3.0 ± 0.9 μm (± SD, n = 45). Note that our FISH samples were not size fractionated, but rather represented "whole" seawater. We also enumerated cells at the surface and DCM for three Florida Straits stations in March, for which the average water sample temperature was 25.0 ± 0.6 °C. On average there were 270 ± 59 cells ml^{-1} at the surface for Stations 01 and 04 with the maximum (312 ± 57 cells ml^{-1}) at Station 04. Cells were also detected at Station 01 and 04 in the DCM, but counts were too low to provide statistically reliable cell concentrations. The same was the case for surface waters at Station 14, however, biliphytes were relatively abundant at DCM (211 ± 39 cells ml^{-1}). It should be noted that FISH-based cell concentrations reflect a conservative estimate since available probes have mismatches to 6 of the 18S rDNA sequences reported herein, including all sequences within BP3.1, and mismatches to 4 sequences from the study by Not, Valentin and colleagues (2007).

Pigment signatures are a common means of analyzing marine FCM samples. We returned to archived FCM data from the same water as collected for FISH and DNA analyses, where greater than 100 biliphytes ml^{-1} were detected (by FISH) to see if a biliphyte population could be identified. The FCM run volumes for these samples was 250 μl, which in theory would yield a cluster of at least 25 cells. While particles with

large (but dissimilar) light scatter relative to 0.75 μm beads, and containing orange and red fluorescence signatures were visible, a coherent population was not. For further verification, an instrument optimized for larger cells and higher sample volume throughput could provide more conclusive results.

Biomass estimate

Based on the hypothesis that biliphytes are photosynthetic, we performed a simple calculation to estimate their contribution to phytoplankton biomass. First, cell volume was estimated using cell size measurements (n = 45) from samples for which we had corresponding Chlorophyll *a* (Chl *a*) data (reflecting total phytoplankton biomass). Given that epifluorescence microscopy only allows measurement of 2 dimensions, a conservative estimate was used for the third dimension by taking the smaller of the two measured dimensions. Therefore, cell biovolume, V, was estimated from the longest cell dimension, L, and the shortest cell dimension, S, both in units of μm using the equation

$$V = (4/3) \times \pi \times ((L/2) \times (S/2) \times (S/2))$$

A conservative carbon conversion factor (2.37×10^{-7} μg C μm^{-3}; Worden et al., 2004) was then combined with the cell abundance and biovolume information. The concentration of total Chl *a* was 0.048 - 0.061 μg Chl *a* l^{-1} (as the sum of monovinyl and divinyl Chl *a*; Hilton et al. in prep.) in the samples (surface March 2005, Stations 01 and 04, Figure 2.1B). These values were similar to previously reported Chl *a* concentrations in the Sargasso Sea (0.051 μg Chl *a* l^{-1}) during a period of stratified oligotrophic waters and over a comparable temperature range (24-27°C), and much lower than concentrations (0.26 to 0.42 μg Chl *a* l^{-1}) in higher nutrient waters (Goericke and Welschmeyer, 1998).

Therefore, biomass of the total phytoplankton community was estimated by converting Chl *a* measurements using a ratio of 90 C:Chl *a*, as appropriate for oligotrophic waters (Eppley, 1968). Biliphytes composed an estimated 28 (±1)% of the total phytoplanktonic carbon in the eddy influenced waters.

Search for a nucleomorph

Should the evolutionary processes leading to extant biliphytes be similar to endosymbiotic events in other protists (see discussion), a nucleomorph could be present. Because of the previous report of a potential nucleomorph-like structure (Not, Valentin et al., 2007), we searched for evidence of a nucleomorph in FISH labeled, DNA (DAPI) counter-stained biliphyte cells. However, DAPI fluorescence was not detected specifically associated with the region of phycobilin-like orange fluorescence (indicating the plastid) that might indicate the presence of a nucleomorph in the FISH illuminated cells. We also accessed approximately 40 *R. salina* cells (an organism known to have a nucleomorph) but could not visualize this structure based on DAPI counter staining.

Because evidence for a nucleomorph was not found in our microscopy work, we tested for the presence of a nucleomorph-like body, of secondary endosymbiotic origin, in biliphytes using a molecular approach. Primers designed to motifs unique to red algae and/or cryptophyte nucleomorph 18S sequences, but biased against the amplification of biliphyte nuclear 18S rRNA genes, were used for PCR and clone library construction. No sequences showing similarity to nucleomorph sequences (or their red algal homologs) were obtained. Of 38 clones for which reliable sequence data was obtained, 33 showed

significant similarity to alveolate and heterokont sequences, with the remaining sequences showing highest similarity to green algae (2), fungi (1), katablepharids (1) and choanoflagellates (1) (data not shown).

Discussion

Evolutionary Relationships

Our phylogenetic analyses revealed a unique eukaryotic lineage (herein named biliphytes) composed of multiple clades with significant statistical support by maximum-likelihood and other phylogenetic methods (Figure 2.2). Clade structure and significance were enhanced by the significant expansion of the number of biliphyte sequences as well as the removal of a likely chimeric sequence (clone HE000427.214, accession DQ222872) used in the previously reported phylogeny (Not, Valentin et al., 2007). In that initial analysis, a clade composed of 3 sequences and considered to be one of 3 overall clades, fell in a basal position. This basal clade is probably an artifact, potentially caused by the inclusion of the chimeric sequence, which falls within it. This artifact probably also destabilized other clades in the former work. At any rate, our analysis did not resolve the basal clade as seen previously (Not, Valentin et al., 2007). Most inner nodes (backbone) did not retain support, as is typically the case for 18S rRNA gene trees, nor did the node which would more firmly establish the sister group to the biliphytes (where bifurcation of the biliphytes from the cryptophytes/katablepharids occurs). This latter node was not bootstrap supported in the analysis of Not, Valentin and colleagues

(2007) either, although a high Bayesian posterior probability value was provided (Not, Valentin et al., 2007). Bayesian posterior probability values are now known to be inflated and should thus be interpreted with caution (Simmons et al., 2004).

18S rRNA phylogenies place the cryptophytes and katablepharids as the closest eukaryotic lineages to the biliphytes, albeit without statistical support (Figure 2.2; Not, Valentin et al., 2007). The cryptophytes (e.g. *R. salina*) acquired photosynthesis secondarily through the uptake and retention of a red algal endosymbiont and are one of only two lineages known to still retain the nucleus of the engulfed eukaryote, i.e., the nucleomorph (Douglas et al., 2001; Palmer, 2003; Archibald, 2007). In the cryptophytes, this DNA containing structure is associated with the plastids, which also contains phycobiliproteins. This result is intriguing given that the FISH and DAPI counterstained biliphyte images of Not, Valentin et al. (2007) indicated the presence of a DAPI-fluorescent body in close association with the plastid, and taken to infer presence of a nucleomorph. We did not observe such a body in association with biliphyte plastids. Our sample resolution is likely highly comparable to those in the former study given that the same microscope model and objective type were employed. However, we did not use an identical counterstaining technique and the fairly high background observed under the DAPI channel could interfere with unambiguous detection of this small structure, which is similar in size to the many marine bacteria that are also stained by DAPI. The findings of Not, Valentin et al. (2007) are reminiscent of the situation in cryptophytes. However, their results do not exclude the possibility that the plastid-localized fluorescence they attributed to nucleomorph DNA is in fact plastid DNA. We were unable to visualize localized DAPI staining in the vicinity of pigment fluorescence in either the biliphyte

cells or in cultures of the cryptophyte *R. salina*, using out whole cell staining methods. Nor could we in a *R. salina* image provided in supplementary materials of the former publication (Not, Valentin et al., 2007). In the case of cryptophytes, an earlier study using embedded and sectioned cells showed that DAPI-stained nucleomorph DNA can indeed be observed, but also revealed a punctate distribution of DAPI-stained plastid DNA (Ludwig and Gibbs, 1987). Thus, at present, it is the presence of a nucleomorph in biliphytes cannot be definitively proven.

If a nucleomorph does exist in the biliphytes, there are (at least) two explanations for its presence. First, the organelle could be homologous to the nucleomorph in cryptophytes, which is reasonable given a possible biliphyte-cryptophyte/katablepharid connection in molecular phylogenies. Under this scenario, the secondary endosymbiosis that gave rise to the plastid and nucleomorph in the two groups would have occurred in their common ancestor. These organelles would then have been completely lost secondarily in katablepharids and the cryptophyte *Goniomonas* (which unlike other cryptophytes does not harbor a nucleomorph). Nucleomorph loss has been proposed for haptophytes, a secondary plastid-containing lineage that large-scale nuclear gene phylogenies suggest are the sister group to cryptophytes (Burki et al., 2007; Patron et al., 2007). Another scenario would be that biliphytes acquired photosynthesis independently through the tertiary engulfment of a cryptophyte or a red alga, which would explain the presence of phycobiliprotein-like fluorescence pigments. Our attempt to amplify nucleomorph 18S ribosomal gene did not reveal any red algal or cryptophyte nucleomorph-like sequences. This suggests that the nucleomorph 18S ribosomal sequence of biliphytes may be quite divergent and, hence, their plastids probably did not

arise through recent uptake of algal cells. However, it is possible that sequencing a greater number of clones might have yielded such a sequence. Alternatively, in contrast to the conclusions of Not, Valentin and colleagues (2007), the biliphytes may not have retained a nucleomorph. Understanding the origin of biliphyte plastids will require the isolation of living cells for ultrastructural, molecular phylogenetic and genomic investigations.

Ecological range and significance in warm waters

Our data offer new insights on the ecological range and genetic diversity of biliphytes. Sequences falling within this lineage were first detected in high latitude marine waters, as cold as -1 °C (Not, Valentin et al., 2007). We show that their range extends through temperate and subtropical oceans, including tropically-influenced features, and throughout the euphotic zone (Figure 2.2). While the Sargasso Sea sites provided excellent representation of subtropical systems, the proximity of the Florida Straits sampling transect to the tropics make this region more representative of tropical waters than common subtropical settings, such as BATS. Typical of tropical waters, such as those that form the Florida Current, high temperatures are maintained year round in the Straits (with extended periods around 30 °C). Furthermore, tropical features can persist (see below) within the Straits. We detected sequences in waters ranging from 5 °C to 30 °C, with nearly one third of the sequences recovered being from waters above 25 °C and sampled over multiple dates (Figure 2.2).

FISH results from the environmental samples support the hypothesis that these organisms are photosynthetic (Not, Valentin et al., 2007), given their phycobilin-like

fluorescence. The abundance of biliphytes in warm water samples was greater than reported elsewhere to date. The only other published biliphyte cell concentrations show that the average (55 ± 23 biliphytes ml^{-1}) for 5 dates at a cold water site in the English Channel, was only a minor portion of the eukaryotic community (on average 1% of total count reported), not including data from 5 dates when none were detected (Not, Valentin et al., 2007). The average temperature at this site for 5 dates spread over the course of the year was 12.5 ± 2.0 °C (D. Vaulot, pers. commun.)

In general, no clear relationship between water temperature and clade structure was visible, suggesting that the organisms within BP1, BP2 and BP3 are able to survive over an extremely broad thermal range (Figure 2.2). Only two supported sub-clades of 18S rRNA gene sequences (BP2.1, BP3.2; Figure 2.2) were composed of sequences from a more narrow temperature range (waters above 20 °C), but this observation should be viewed with caution until more data are available. 18S rRNA gene data initially revealed the "picobiliphytes" in cold surface waters of the Arctic Ocean, the Norwegian Sea, and in some fall and winter coastal European waters (Romari and Vaulot, 2004; Not, Valentin et al., 2007) as well as a single sequence from the permanent thermocline (500 m) near BATS (Not, Valentin et al., 2007). The fact that a sequence (SSRPB47, accession EF172850) was retrieved from 500 m could have several inferences, e.g.: 1) that primary production by these organisms can be exported to the deep sea, serving to sequester atmospheric CO_2; or 2) that these organisms are mixotrophic or even heterotrophic (although heterotrophy seems unlikely). Although only a single sequence was recovered, the finding at 500 m is somewhat unusual for phytoplankton, as algal sequences are rare in the few size fractionated 18S rRNA gene libraries from the permanent thermocline or

below (López-García et al., 2001; Not, Valentin et al., 2007). Sequences falling within the biliphytes, although not identified as such when published, were also obtained from an environmental clone library in 23^{o}C summer-time surface coastal waters (Countway et al., 2005).

Biliphyte sequences were not found within two DCM libraries from sites where biliphyte sequences were recovered at the surface. For example, the March Florida Straits Station 04 DCM library, was unlike most other Florida Straits libraries as a large number of prasinophyte clones were sequenced (*Bathycoccus*, *Micromonas* and *Ostreococcus* composed one third of those sequenced, with the majority of others belonging to the novel alveolates, as commonly found (Groisillier et al., 2006; Worden, 2006; Not et al., 2007, data not shown) but no biliphytes sequences were recovered. Similar to Station 04 results, the Station 01 DCM library contained many prasinophytes in March (data not shown). Nevertheless, biliphytes were detected by FISH in both these samples, although at extremely low abundance (see results). The third library from which biliphyte sequences were not retrieved was the Northern Sargasso Sea DCM library. Here again, and unlike other Sargasso Sea samples, a large number of prasinophyte sequences were retrieved (data not shown).

It is noteworthy that during the March 2005 cruise, conditions in the western half of the sampling transect were strongly influenced by remnants of a mesoscale cyclonic eddy. Continuous surface measurements of temperature, salinity and fluorescence recorded underway indicate a distinct front was present on 31 March 2005 at around 79.85 W. Colder, more saline (data not shown) euphotic zone waters with higher fluorescence occurred on the Florida side of this front (Figure 2.1D). Satellite imagery

from around this time (imars.usf.edu) indicates these characteristics were associated with upwelled waters of a mesoscale eddy translating downstream. Previous work has shown that these cyclonic features, termed Tortugas Eddies, form as frontal eddies along the tropical Loop Current, translate into the southern portion of the Florida Straits and remain resident there for a period of time before moving downstream where they elongate and shear apart off of the upper Florida Keys (Lee et al., 1994; Fratantoni et al., 1998). The shoaling of the thermocline in cyclonic eddies can increase nutrients supply to the euphotic zone and ultimately primary production (McGillicuddy et al., 2007). Enumeration of biliphytes by FISH for the March cruise reveals highest biliphyte abundances at the surface of Station 01 and 04, both located on the Florida side of the front, where tropical eddy waters had an influential role on the water mass.

Biliphytes were abundant in the surface waters of Station 01 and 04 (the eddy influences waters) and DCM of Station 14 in March, although low in abundance in other samples from the same cruise. Although phylogenetic analysis showed no clear relationship between clade structure and water temperature, the fact that waters with higher abundances were isothermal (24 °C; Figure 2.1D, Table 2.1) suggests a potential optima related to temperature or associated environmental parameters. Further analysis and enumeration of biliphytes and other taxa in a range of conditions is clearly needed. The higher abundances were found in both low light (DCM) and high light (surface) isothermal waters. These waters (west of the front, influenced by the eddy; and DCM waters east of the front, influenced by higher nutrient concentrations found at the base of the euphotic zone) likely had relatively high nutrient concentrations. With respect to niche differentiation, it is then possible that conditions during which prasinophytes

proliferated (particularly the significantly lower temperature DCM waters (Figure 2.1D) at Florida Straits Stations 01 and 04 in March, data not shown) are less advantageous to biliphytes. The eddy waters clearly contributed to a shoaling of the thermocline at these stations, likely yielding higher nutrient in the warm surface waters where the biliphytes appeared to thrive.

Cell size, enumeration and biomass

The nanoplanktonic size of these organisms seems at odds with the designation "picobiliphytes" given by Not, Valentin and colleagues (2007). Whether the denomination "pico", refers to cells <2 μm in diameter (Sieburth et al., 1978; Stockner, 1998; Li et al., 2006), or, as defined by a subset of studies, to cells <3 μm diameter (Romari and Vaulot, 2004; Not, Valentin et al., 2007), it is evident that a large portion of the probed cells are greater in diameter than indicated by either of these definitions. Images in the supplementary materials of the former publication (Not, Valentin et al., 2007) show several relatively spherical cells (akin to ours) which, based on the scale bar provided, are approximately 5 x 5 μm, different than the shape and size estimate provided within the text (2 x 6 μm cells, n=9). Together with our cell size data, this makes the denomination "pico-" a confusing prefix. Competition processes, including those related to optimization of surface area to volume ratios, photosynthetic properties and grazing are all critical factors for population dynamics and are linked to organism size. The fact that sequences from larger cells are retrieved from our size fractionated (<2 μm) DNA samples is not necessarily surprising, as fragile cells and even appendages are known to break during the pre-filtration process, resulting in presence of some sequence data from

larger organisms (see Worden, 2006; Not, Valentin et al., 2007). At the same time, no biliphyte sequences were identified in the GOS data (Rusch et al., 2007), which used a <0.8 μm size fractionation step. Sequences from known picoplanktonic organisms, larger than this cutoff, in particular the prasinophyte *Micromonas pusilla*, which averages from 1.4 -1.6 μm in diameter based on coulter counter data from cultured strains (Welsh et al. in prep.), are present at several GOS sites, showing that at least small phytoplankton did pass through the pre-filter. The absence of biliphytes could be based either on the larger cell size of these organisms leading to complete exclusion, or its relative rareness, making the sequencing coverage inadequate for recovering an 18S gene fragment amongst genetic material from many more abundant organisms. Given the importance of ecological inferences associated with marine plankton size classes, and given that there are many plankton species falling exclusively within the <2 μm size fraction, precise size classification is a fundamental ecological distinction. Hence, while lack of cultured isolates hinders further characterization, it seems prudent that this novel group be named "biliphytes", in keeping with (Not, Valentin et al., 2007), but removing the prefix referring to size classification.

Our data indicates that reevaluation of some previous studies on phytoplankton dynamics and biomass contributions will be needed. Cryptophytes have been reported in FCM and high performance liquid chromatography (HPLC) studies, but few cryptophyte-like sequences appear in open-ocean clone libraries (Countway et al., 2007; Not et al., 2007). In the case of FCM, the similar orange fluorescence between cryptophytes and biliphytes, may have caused biliphytes to mistakenly be identified as cryptophytes in earlier studies, depending on factors such as suitability of FCM fixation methods for

biliphyte preservation. FCM fixation and cryo-freezing procedures can be disruptive, particularly for eukaryotic nanoplankton. We quantified cell loss and found that while cyanobacterial numbers quantified in either glutaraldehyde (0.25%, final concentration) or paraformaldehyde (1%, final concentration) cryo-frozen samples were 100% of live counts, eukaryotes generally suffer some losses. This was documented using mid-exponential growth cultures for several phytoplankton strains. The prasinophytes *Ostreococcus tauri* and *Micromonas pusilla* as well as the prymnesiophyte *Isochrysis* sp. CCMP1244 preserved relatively well, with losses ranging from 0% to 22%, depending on the cell type and fixative used (fixative always used in conjunction with flash freezing in liquid nitrogen). However, the cryptophyte *Rhodomonas salina*, a cell wall-less nanoplankter, suffered 40 ± 7% loss in glutaraldehyde and 93 ± 3% loss in paraformaldehyde (Welsh et al. in prep.), both with cryo-freezing. Similar observations have previously been reported for *R. salina*, using a range of fixatives although without cryo-freezing (Klein Breteler, 1985). Because quantification of fixation effects cannot be performed without cultures, we can only infer, based on our FCM data and any potential similarities to species such as *R. salina*, that at least some fraction of biliphytes are likely lost via many of the common oceanographic FCM preservation methods. This could result in both a lack of an identifiable biliphyte population and an underestimation of the total eukaryotic phytoplankton community.

With respect to HPLC, these analyses do not utilize phycobiliproteins to quantify specific groups and therefore the presence of phycobilins in both cryptophytes and biliphytes would not confound interpretation of previous HPLC studies. However, currently it is unclear whether the signature pigments (Latasa et al., 2001; Latasa, 2007)

used for cryptophytes (alloxanthin) might also be present in biliphytes; this would depend on the nature of the secondary endosymbiosis event and downstream processes.

Given their large size relative to the numerically dominant phytoplankton in oligotrophic environments (e.g. picoplankton such as *Prochlorococcus*), low numbers of these putative primary producers can contribute disproportionately to phytoplankton biomass. In the Florida Straits, the picocyanobacteria *Prochlorococcus* was always responsible for >30% of the phytoplankton biomass, as might be expected for warm stratified oligotrophic environments (Hilton et al. in prep.). Nevertheless, biliphytes contributed an estimated 28% of the total phytoplanktonic carbon in the eddy influenced, surface waters. This is a large fraction for a single taxon, and could be related to eddy dynamics, which often enhance primary production (McGillicuddy et al., 2007). Moreover, the impact of biliphytes may be greater in warm waters than in cold waters due to the higher abundances. High latitude sites sampled within the previous work (Not, Valentin et al., 2007) are influenced by Atlantic Ocean currents so that the identified populations could reflect residual populations from warmer waters. Clearly, much remains to be learned about the abundance, dynamics and controls of biliphytes. Their widespread distribution is remarkable and presumably reflects a high capacity for acclimation to dramatic temperature variations.

Conclusions

The discovery of biliphytes at a well-studied time-series site (i.e., BATS) reflects the current paucity of our knowledge concerning eukaryotic plankton; unless these novel eukaryotes are indeed a recent addition to the community. While the relative importance

of such novel microbes to primary production remains to be determined, their discovery necessitates reconsideration of phytoplankton dynamics. Our data from warm subtropical and cool temperate environments, combined with previously reported sequences from cold high latitude waters (Not, Valentin et al., 2007), underscores the need to define and quantify the role of biliphytes in primary production and carbon cycling. Dynamics of phytoplankton with broad thermal ranges may be particularly important when considering future climatic scenarios.

Acknowledgements

Alejandra Ortiz, Eunsoo Kim, Heike Moehlig, David E. Richardson, John F. Heidelberg, John M. Archibald, and Alexandra Z. Worden contributed to this work and are listed as co-authors on the peer-reviewed publication based on this dissertation chapter. R. Gausling, R. Welsh and K. Penn assisted with clone libraries; R. Cowen and B. J. Binder provided cruise space and the work was facilitated by the captains and crews of the R/V Oceanus, R/V Endeavor, R/V Walton Smith; C. Guigand facilitated cruises and developing figures; A. Engman, J. A. Hilton, F. Not, R. Welsh and L. Zamora helped with cruises; M. Latasa performed HPLC analysis; M. Everett ran Model Test; S. McDonald and especially D. Hansell provided helpful discussions and comments. Aspects of the work were supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada to J. Archibald and the primary funds were through an Investigator Award from the Gordon and Betty Moore Foundation and an NSF grant OCE-0623928, both to A.Z. Worden.

Chapter 3:

Ecology of uncultured prymnesiophytes

Summary

Oceanic primary production accounts for half the global carbon dioxide fixed annually. A diverse assemblage of phytoplankton conducts marine photosynthesis yet the ecological roles and physiology of the eukaryotic component is poorly understood. Complete genomes have been analyzed for only two marine eukaryotic phytoplankton lineages, prasinophytes and stramenopiles, although tiny 'picoplanktonic' members of a third lineage, prymnesiophytes, have long been inferred to be of oceanographic importance. We investigated the ecological importance of natural pico-prymnesiophyte populations using cultivation-independent methods. Pico-prymnesiophytes flow sorted from subtropical waters belonged to broadly distributed but uncultivated taxa. Field experiments showed that uncultivated pico-prymnesiophytes grow rapidly and contributed significantly to primary production. They also formed a large portion of global picophytoplankton biomass, with differing contributions in five biogeographical provinces, from tropical to high latitudes.

Background

Global primary production is partitioned equally among terrestrial and marine ecosystems each accounting for approximately 50 gigatons (10^9 tons) of carbon per year (Field et al., 1998). The photosynthetic plankton mediating marine primary production are derived from an array of evolutionary histories. In the open ocean, they include

cyanobacteria (e.g. *Prochlorococcus* and *Synechococcus*) and a diverse set of eukaryotic phytoplankton such as diatoms, prasinophytes and prymnesiophytes, the latter also being referred to as haptophytes (Chisholm et al., 1992; Keeling et al., 2005; Hampl et al., 2009). Most oceanic phytoplankton are very small or 'picoplanktonic' (<2-3 μm diameter), which is an advantage in the low-nutrient conditions frequently encountered in open ocean, due to enhancement of resource uptake via greater surface area to volume ratios (Chisholm, 1992; Raven, 1998; Vaulot et al., 2008; Worden and Not, 2008). In contrast, many episodic bloomers like diatoms, dinoflagellates and some prymnesiophytes, including the coccolithophore *E. huxleyi*, are larger in size. These larger cells were considered responsible for most carbon export to the deep ocean, however picophytoplankton can also play substantial roles in carbon export (Richardson and Jackson, 2007). Despite their importance to carbon cycling, genomic information on eukaryotic phytoplankton is still sparse. The six eukaryotic phytoplankton genomes that have been completely sequenced and analyzed comparatively show greater levels of genomic differentiation than originally anticipated based on the small subunit (SSU) ribosomal RNA (rRNA) gene (Palenik et al., 2007; Bowler et al., 2008; Worden et al., 2009). The genomic divergence observed translates to major differences in physiology and adaptation to an ecological niche (Parker et al., 2008; Bowler et al., 2010).

Pigment based estimates indicate prymnesiophytes are broadly distributed and abundant phytoplankton. Oceanic prymnesiophytes are thought to be small based on high levels of prymnesiophyte-indicative pigments in regions where most chlorophyll *a* (Chl *a*), representing the total phytoplankton community, is in the <2 μm size fraction (e.g. Bidigare and Ondrusek, 1996; Mackey et al., 2002; Worden and Not, 2008). Six

picoplanktonic prymnesiophytes exist in culture (Vaulot et al., 2008; Worden and Not, 2008). However, prymnesiophyte sequences from <2-3 μm size fractioned environmental rDNA clone libraries typically belong to uncultured taxa (Moon-van der Staay et al., 2000; Not et al., 2007; Worden and Not, 2008, Liu et al., 2009). As a whole, this lineage evolved early in the history of eukaryotic phytoplankton as the result of a secondary endosymbiosis event and their overall placement in the eukaryotic tree of life is still uncertain (Keeling et al., 2005; Medlin et al., 2008). Prymnesiophytes are estimated to have diverged from other major eukaryotic lineages 1.2 billion years ago (Medlin et al., 2008), making them extremely distant from phytoplankton with published genome sequences, such as diatoms (belonging to the stramenopiles) and prasinophytes (Archaeplastida). Thus, while many inferences exist regarding oceanographic importance and evolutionary history, basic uncertainties surround cell size, biomass, environmental growth rates and genomic composition of oceanic prymnesiophytes. We undertook parallel, cultivation-independent approaches to analyze pico-prymnesiophytes by comparative genomics and to establish their environmental contributions, including specific growth rates.

Materials and Methods

Field sampling

Eleven oceanographic expeditions were performed. Samples for enumeration of the four picophytoplankton groups were collected on transect cruises in various ocean basins: N92S (equatorial Pacific Ocean, 25 April-5 May 1992), N92F (equatorial Pacific Ocean, 10-17 September 1992), N93 (Atlantic Ocean, 7 July-28 August 1993), N95 (Indian

Ocean, 24 September-23 October 1995), N96 (SW Pacific Ocean/Southern Ocean, 19 January-1 February 1996), S201 (NE Pacific Ocean, 14-25 March 2001) and a transect from coastal New England to the Bermuda Atlantic Time-series Study (BATS) station (OC413, 23 May-12 June 2005). Sampling was more intensive for this latter cruise. Higher resolution flow cytometry (FCM, 12 depths) and microcopy sampling were collected than for the above cruises. In addition, DNA samples (surface and deep chlorophyll maximum, DCM, as determined by *in situ* fluorometry) were collected and dilution experiments performed according to the methods of Landry and Hassett (1982) using modifications in Worden and Binder (2003). The experimental work was conducted at two main stations: BATS, which had summer-like conditions with the mixed layer extending to 30 m and a Northern Sargasso Sea (NSS) station (35° 10′ N, 66° 33′ W) where summer-like conditions were less pronounced and the mixed layer extended slightly deeper, to 40 m. FCM, microscopy and DNA samples were prepared from both standard water column samples as well as dilution experiments. In addition to these 7 expeditions more in depth seasonal sampling was performed across three Gulf Stream stations in the Florida Straits between South Florida, USA and the Bahamas in 2005 (cruises WS0503, 31 March-1 April; WS0510, 18-19 May; WS0515, 24-25 June; WS0518, 31 July-1 August; WS0523, 27-28 September; WS0528, 7-8 December). For these cruises FCM, microcopy, DNA and high performance liquid chromatography (HPLC) samples were collected at the surface and DCM. In 2007, sampling and flow sorting was performed along this same transect (WS0705, 27 February 2007). For three additional cruises only DNA samples were analyzed (surface and DCM); these were in the Western Pacific (CN207, October 2007) and the North Atlantic Ocean (EN351 and

EN360 in April and September 2001). For all the cruises, seawater was collected using either GO-FLO bottles or a Sea-Bird Niskin Rosette equipped with standard CTD and photosynthetically active radiation (PAR) detectors. Vertical profiles of temperature, salinity, irradiance and fluorometry were recorded *in situ* at each station.

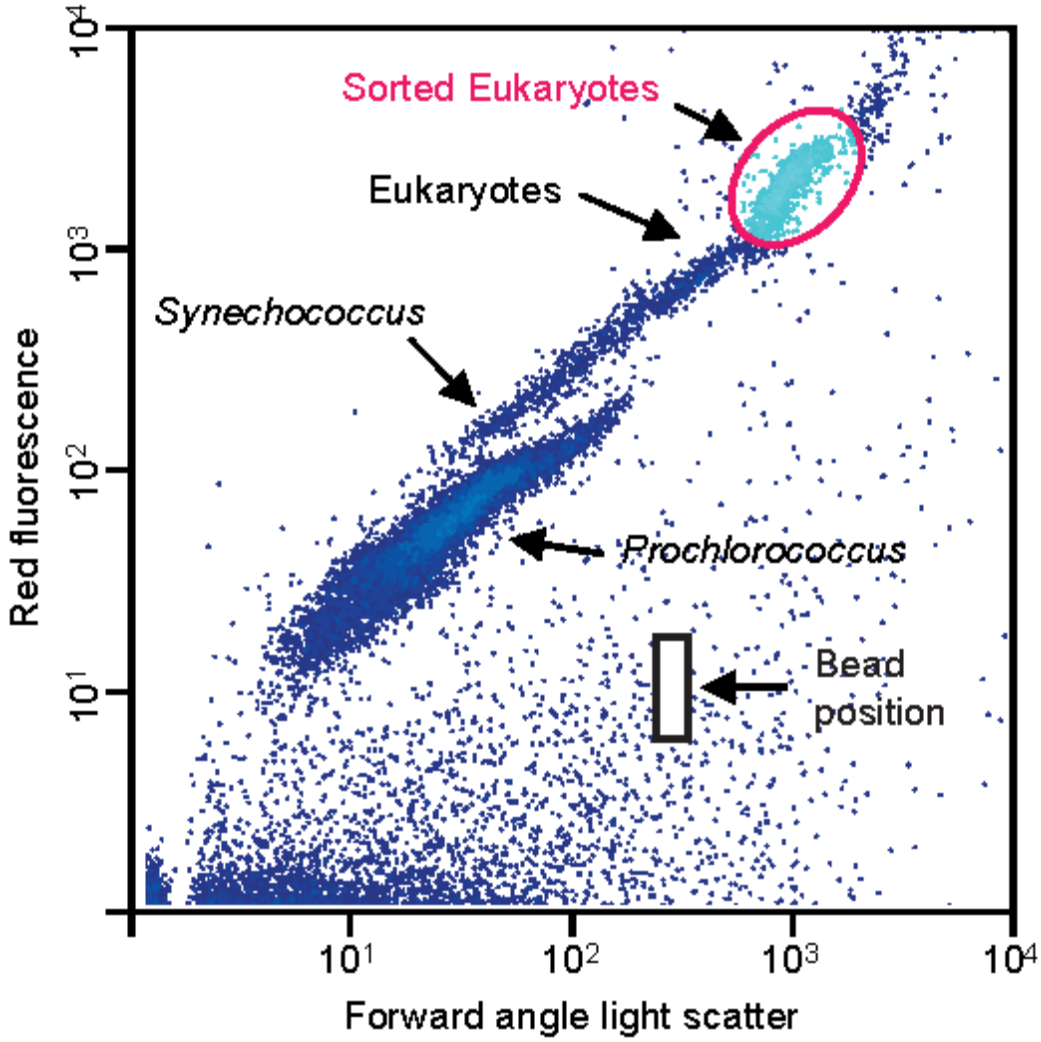
Flow cytometry and cell sorting

The instrument used for flow sorting (Influx, Cytopeia Inc., now Becton Dickinson, Franklin Lakes, NJ, USA) was equipped with a 488 nm laser (200 mW output), a 70 µm diameter nozzle and generally run at a flow rate of 25 µl min^{-1}. Forward angle light scatter (FALS), pulse width, side scatter (90° angle; SSC), red autofluorescence from chlorophyll (692 ± 40 nm band-pass filter) and orange autofluorescence from phycobiliproteins (527 ± 27 nm band-pass filter) were used as population discriminators and also recorded (log-integrated for scatter and fluorescence). The trigger was FALS. This instrument is only ever run using sterile solutions as sheath and is always cleaned extensively and air dried prior to shutdown. An iterative approach was taken to 'learning' which populations seen during sample analysis corresponded to particular taxa, or groups of taxa. This was done by performing sorts and characterizing them using 18S and 16S rRNA gene libraries after performing multiple displacement amplification (MDA) on the samples (see later methods sections), allowing us here to specifically target the population we knew to contain pico-prymnesiophytes. For all of these sorts, the instrument was also cleaned extensively upon set-up to eliminate potential sources of contamination. The sheath and sample lines were sterilized prior to running samples using a series of 10% bleach (in 0.2 µm filtered 18.2 MΩ water); 0.2 µm filtered 18.2

MΩ water; and finally 70% ethanol (in 0.2 µm filtered 18.2 MΩ water). Specifically, the sheath tank was filled with the 10% bleach solution which was run through all the lines (sheath and sample) for 5 min at a high flow rate. While the lines were filled with the bleach solution, a series of on/off cycles were performed for all pneumatic valves to disrupt any possible particulate build up. The sheath tank was then removed, emptied and rinsed 5 times, filled with 0.2 µm filtered 18.2 MΩ water, and the lines then rinsed for 10 min. Finally, the sheath tank was filled with the 70% ethanol solution and run just long enough to fill all the sample and sheath lines before stopping the run; all lines were then 'blown dry' with compressed air. In all sorts, 1X phosphate buffer saline (PBS) solution was used as sheath fluid. PBS was prepared from a sterile 20X solution with 0.2 µm filtered 18.2 MΩ water and was subsequently 0.2 µm filter sterilized again and autoclaved prior to use. Seawater samples were sorted within hours of collection.

Control sorts (to test for contamination) included sheath fluid collected from the sheath reservoir as well as collected from the test stream which run through the sample lines (see below). The piezo amplitude was 0.56 Volts, drop frequency of 49.3 kHz and a drop delay approximately 34.5 droplets. Populations were selected based on specific SSC, FALS, pulse width and chlorophyll autofluorescence criteria with gates from all of these parameters used to define each of the single populations sorted (Figure 3.1), increasing the stringency, although also decreasing yields as only particles meeting all criteria were sorted. Approximately 300 cells were sorted (determined by a post-analysis rerun of sorted sample replicates).

Figure 3.1. Forward angle light scatter and chlorophyll-derived red fluorescence characteristics of picophytoplankton groups in the Gulf Stream core (blue arrow, inset Figure 3.2). Targeted metagenome DNA came from the population labeled "sorted eukaryotes" (aqua, magenta circle). The box (black) shows the position of 0.75 µm bead standards run in a previous sample, under the same settings, but not added to sorted samples to avoid contamination. Axes are shown in log-scale.



Flow cytometry was also used to determine picophytoplankton cell counts on the global transects. For “WS”, and “OC” cruises FCM samples (1 ml) were collected in triplicate for each depth, preserved with 0.25% glutaraldehyde (final concentration; Tousimis, Rockville, MD, USA) and stored in the dark for 20 min, a modification of previously published methods (Olson et al., 1993). Again, data collection was triggered on FALS and instrument setup/data collection was as above (InFlux). The samples were then flash frozen in liquid nitrogen and either stored in liquid nitrogen until processing or kept at -80 °C for long-term storage. Samples were thawed just prior to analysis and a known volume of yellow-green 0.75 µm beads (Polysciences, Inc., Warrington, PA, USA) was added and used as internal fluorescence and light scatter standards (Olson et al., 1993).

For all “N” and “S” cruises, samples (1.8 ml) were fixed in 0.2% paraformaldehyde (PFA; final concentration), stored in liquid nitrogen. Samples were analyzed using a Coulter EPICS 753 flow cytometer equipped with two 5 W Aragon lasers and data was collected for picoplankton abundance and red fluorescence per cell according to previously published protocols (Campbell and Vaulot, 1993; Monger and Landry, 1993; Campbell et al., 1994; Buck et al., 1996).

Listmodes were analyzed either using CYTOPC software, see Campbell and Vaulot (1993), or WinList (Verity Software House, Topsham, ME, USA). *Prochlorococcus* and *Synechococcus* were defined based on their FALS and fluorescence characteristics (Olson et al., 1990), with the PMTs at relatively high voltage settings in order to simultaneously enumerate *Prochlorococcus*. Note that for a small number of samples (indicated where appropriate) the *Prochlorococcus* intersected with baseline noise in the red-fluorescence

channel. These samples were not used for biomass averages (Supplementary Information Section II) or for pico-prymnesiophyte contributions to total phytoplankton biomass (rather the data for all groups was 'thrown away'). 'Non-prymnesiophyte' picoeukaryotes were enumerated using the same analysis windows as by Buck and colleagues, who showed that for field samples analysis using these windows (encircling the smallest eukaryotes) rendered FCM results that were tightly correlated with the sum of all picoeukaryotes, excluding pico-prymnesiophytes, enumerated by microscopy (Buck et al., 1996).

Whole genome amplification and sample prescreening

Natural populations (and controls) were flow sorted with capture in two directions, right and left, using the InFlux as described above, directly into nuclease/pyrogen free cryovials and frozen at -80 °C. The volume of the sorted droplet was ~1 µl. A subset of replicate sorts were immediately resuspended and rerun to determine the sort efficiency (32% for right sorts, used for metagenomic analysis), i.e., the true number of cells that would then go into the MDA reaction used to amplify the whole genome, detailed below. For whole genome amplification, we also tested the efficiency of the lysis used (KOH, 10 min, on ice) showing that ~54% of the sorted cells for the right sorts used herein were lysed (data not shown). The following environmental samples underwent MDA (Repli-g Midi kit, Qiagen, Germantown, MD, USA) post the alkaline lysis step (all in duplicate): right sort population (Station 04, 75 m; used for metagenomic analysis), left sort population (Station 04, 75 m), right sort population (Station 08, 141 m). Duplicates of each of three different control sorts also underwent MDA, specifically: 1) sheath run

through entire sheath system, collected after sort test deflection, prior to introduction of seawater samples to the system on the sort day; 2) sheath fluid run as sample through the sample line collected in tube using sort test deflection; 3) sheath run through entire sheath system, collected after sort test deflection (but performed later in the day, after environmental sorts). Hence, environmental sorts were performed between controls 2 and 3. Finally, a positive DNA control (100 pg gDNA, Qiagen, Germantown, MD, USA) and a negative control (H_2O) also underwent MDA (both in duplicate). After storage at -80 °C, sort samples and the above controls were transferred to a thin-walled microfuge tube, sample volume determined and brought to a total volume of 2 µl with TE-buffer; negative controls were performed using the same 'template' volumes. The total reaction volume was 10 µl and all handling of reagents and samples were performed in a polymerase chain reaction (PCR)-workstation with HEPA-filtered air supply. The reactions were incubated at 30 °C for 16 h according the recommended protocol from the manufacturer (Repli-g Handbook, Qiagen, Germantown, MD, USA). The amplifications were subsequently diluted 5-fold with TE-buffer before heat-inactivation at 65 °C for 3 min. The MDA products then served as template for PCR to construct preliminary 16S and 18S rRNA gene libraries (see below) to select a sample to advance for metagenomic sequencing. The purpose of this quality control step was also to verify potential contamination in the sample handling from collection through whole genome amplification. Negative controls (H_2O and all sheath sorts) did not render 16S or 18S rRNA gene PCR products. A small number of clones from the flow sorted phytoplankton population were then sequenced (~10 clones per replicate) and used to screen different MDA products for target organisms (from right sorts at Station 04 and Station 08).

Size fractionated 18S rRNA gene clone libraries

Standard clone libraries were generated for multiple samples from three Florida Straits cruises, WS0503, WS0518, and WS0528 (Station 14 only) and all the Sargasso Sea cruises (EN360, EN351 and OC413), as well as the western Pacific cruise CN207. Environmental conditions and clone library specifics are shown in (Supplementary Information Section I Table S3.1). In general, 1 l of seawater (0.5 l for OC413) was collected from the surface and DCM for all the cruises except for OC413 for which samples were collected at 15 m and 75 m for the BATS station and at 15 m and 70 m for the NSS station. Seawater was pre-filtered (for size fractionation) by gravity though a 2 µm polycarbonate filter (GE Osmonics, Minnetonka, MN, USA). The water was then vacuum filtered onto a 0.2 µm (OC413, WS0503, WS0518, WS0528; <10 mm Hg) or 0.45 µm (EN351, EN360; <5 mm Hg) Supor filter (Pall Gelman, Ann Arbor, MI, USA). The Supor filter was immediately frozen cryogenically and subsequently moved to -80 °C for long term storage until further analysis. For CN207, seawater from the Niskin rosette was transported to a large reservoir that had been cleaned with a 10% HCl solution. Cells were collected on a 0.8 µm pore size, 293 mm Supor filter (Pall Gelman, Ann Arbor, MI, USA) after pre-filtration through a 3 µm pore size filter (in series, both under vacuum). Prior to collecting samples the entire filtration system and reservoirs were flushed with a solution composed of 1:9 bleach:18.2 MΩ H_2O to reduce the possibility of contamination. These large filters were flash frozen by suspension in liquid nitrogen vapor and stored at -40 °C.

For all but CN207 clone libraries, DNA was extracted using the DNeasy kit (Qiagen, Germantown, MD, USA) according to the manufacturer instructions. For CN207, a sucrose extraction protocol (http://www.mbari.org/phyto-genome/Resources.html) was used to extract DNA from the 293 mm filters.

18S rRNA genes were amplified using primers (Forward 5′-ACCTGGTTGATCCTGCCAG-3′ and Reverse 5′-TGATCCTTCYGCAGGTTCAC-3′) complimentary to conserved regions proximal to the gene termini, designed as universal eukaryotic primer set, but likely with some biases (Moon-van der Staay et al., 2000; Moon-van der Staay et al., 2001). In both cases, PCR was performed; with an initial 'hot start' for 15 min at 95 °C, proceeded by 32 cycles at 94 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min; followed by a final extension at 72 °C for 10 min, as in Cuvelier et al., (2008). One µl of PCR product was ligated into the vector pCR2.1 using the TOPO-TA cloning kit (Invitrogen, Carlsbad, CA, USA) and transformed, after colony selection and growth plasmids were purified. "WS" and "OC" cruises were sequenced with a single primer internal to the PCR product (502F) that rendered a unidirectional product for all 18S rRNA gene clones. "CN" and "EN" cruises were sequenced with a suite of primers, two plasmid targeted primers (M13F and M13R) and two primers internal to the product (1174R and 502F). Sequencing was performed using the Big Dye Terminator chemistry on an AB3730xl sequencers (Applied Biosystems, Foster City, CA, USA). For "WS" and "OC" cruises, 96 clones were sequenced per library, for "EN" cruises 40 clones were sequenced per library.

BLASTN against the GenBank non-redundant (nr) database was used to make a preliminary taxonomic affiliation for the sequences obtained from the clone libraries. In

the Sargasso Sea, prymnesiophyte sequences were retrieved from all four environmental clone libraries (BATS and NSS stations at the surface and DCM) and accounted for 1 to 12% of the total numbers (96) of sequences in each library. In the Florida Straits, prymnesiophyte sequences were retrieved from 13 of the 14 clone libraries (1 to 17% of the total number of sequences in each library). Chromatograms and assemblies of all 18S rRNA gene sequences tentatively assigned to the prymnesiophytes were manually curated. For the phylogenetic analysis, only the curated sequences were analyzed, alongside prymnesiophyte sequences retrieved from GenBank as well as a suite of outgroup sequences (using BLASTN, last retrieval February 2009). Manual screening was use to detect for chimeras which were subsequently removed. Sequences were aligned with those retrieved from GenBank using ClustalW. Preliminary neighbor-joining trees were built using modules within the PHYLIP package (Felsenstein, 2005) and 280 sequences (including outgroups). Generally only a single representative of cultured species, or strain, as well as a single representative from each clone library found within a single clade, was kept for subsequent phylogenetic analysis. 139 environmental sequences were then used in the final tree, including 111 from our samples (72 from the Florida Straits, 27 from the Sargasso Sea, 7 from the Western Pacific and 5 from the MDA, see below). We also used 28 environmental sequences retrieved from GenBank, 5 from an earlier study of ours in the Sargasso Sea (Not et al., 2007), 4 from the Indian Ocean (Not et al., 2008), 13 from the equatorial Pacific (Moon-van der Staay et al., 2000; Moon-van der Staay et al., 2001), 2 from L'Atalante deep-sea basin (Alexander et al., 2009), 1 from coastal subtropical western Pacific (Cheung et al., 2008), 1 from the North western coastal Mediterranean Sea (Massana et al., 2004a), 1 from the Southern

California Bight (Worden, 2006) and 1 from unknown origins (location is not specified in the GenBank entry). This represented all environmental prymnesiophyte 18S rRNA gene sequences housed at GenBank as of February, 2009, that has sufficient overlap with our sequenced products to be included alignments. These sequences were then realigned with ClustalW and the alignment was manually edited. The final overview 18S phylogenetic analysis was performed using maximum likelihood methods (PhyML, Guindon and Gascuel, 2003) after prediction of the best evolutionary model (in this case GTR+I+G) using ModelTest (Posada and Crandall, 1998). Model parameters used were 1.5886, T_iT_v; 0.3044, p_{inv}; 0.6026, gamma distribution shape (alpha). Data was bootstrapped with 100 replicates. A number of 18S rDNA sequences from red and green lineage organisms served as outgroups, these were: *Chondrus cripus*, *Gracilaria lemaneiformi*, *Compsopogon coeruleus*, *Cryptomonas ovate*, *Cryptomonas pyrenoidifera, Hemiselmis virescens, Rhodomonas salina, Pyrenomonas helgolandii, Pyramimonas australis, Chlamydomonas reinhardtii, Chlorella vulgaris, Micromonas pusilla strain CCMP1723, Micromonas pusilla strain CCMP1545, Symbiodinium microadriaticum, Prorocentrum micans, Karlodinium micrum, Coscinodiscus radiates, Thalassiosira weissflogii, Thalassiosira pseudonana, Gloeochaete wittrockiana, Cyanophora paradoxa, Glaucocystis nostochinearum.*

Prymnesiophyte cell counts

Prymnesiophytes were identified by fluorescence *in situ* hybridization (FISH) using a prymnesiophyte-specific probe (Not et al., 2002) or using a characteristics-based method based on their chloroplast arrangements, flagellar characteristics and occasionally the

presence of a haptonema (Andersen, 2004). No significant difference (t-test, $p = 0.428$) was detected between these two microscopy methods. Comparison of data from between 25° to 35° in the Atlantic showed comparably abundances, with the prymnesiophyte characteristics-based average being 593 ± 108 ml^{-1} (SE, n=12) while the FISH average was 500 ± 61 ml^{-1} (SE, n=26) for different sample sets, and cruises.

FISH was performed on "WS" and "OC" cruise samples, using a prymnesiophyte specific probe and probed to enumerate cells by epifluorescence microscopy. To prepare and store samples for hybridization, seawater (180 ml for OC413 CTD profiles and 100% raw seawater treatments in the dilution experiments; 405 ml of seawater for the 40% and 20% raw seawater; 90 ml for all "WS" 2005 cruises) was preserved with PFA (1%, final concentration) for a minimum of 1 hour at 4 °C in the dark. The seawater was filtered onto a 0.2 μm Anodisc (25 mm, Whatman, Maidstone, UK), the filters were dried with an ethanol series (50%, 80% and 100% ethanol diluted in autoclaved 18.2 MΩ water for 3 min each) and stored at -80 °C prior to hybridization. FISH was performed on replicate filter pieces in conjunction with tyramide signal amplification (TSA) using a modification of a previously published method (Not et al., 2002; Cuvelier et al., 2008). We used an oligonucleotide probe specific for prymnesiophytes (PRYM02, 5′ GGA ATA CGA GTG CCC CTG AC 3′, Simon et al., 2000). This probe has no mismatches with the prymnesiophyte 18S rRNA gene sequences in the main tree (Figure 3.2), with the following exceptions: OLI16029, 1 mismatch; OLI51033, 2 mismatches; OLI51059, 2 mismatches; OC413BATS_O071_75m, 2 mismatches; FS01AA77_01Aug05_5m, 1 mismatch; *Chrysochromulina leadbeateri*, 3 mismatches; *Chrysoculter rhomboideus*, 1 mismatch. Note that several of the OLI sequences had gaps, or apparent nucleotide

substitutions, in several highly conserved positions for other eukaryotes. Hybridization efficiency of PRYM02 was tested on a culture of a larger cultured prymnesiophyte species, *Isochrysis sp.* CCMP1244; out of the 1,492 cells detected using the DNA specific dye 4',6-diamidino-2-phenylindole (DAPI), 1,480 cells (or 99.3 ± 3.3% of the cells) were positively hybridized. After hybridization according to methods in Not et al., (2002) and Cuvelier et al. (2008) FISH filters were counter-stained with DAPI. This was performed by counterstaining with 2.5 μg ml^{-1} for 5 min, rinsing for 5 min at room temperature (RT) in autoclaved 18.2 MΩ water, briefly dipping in 80% ethanol and then air drying for about 10 min, and finally applying 7 μl of mounting solution [1:5 antifading solutionAF1 (Citifluor, London, UK) and Vectashield mounting medium (Vector Laboratory, Burlingame, CA, USA) for "OC" samples. For "WS" samples, filters were air dried for about 15 min and Vectashield mounting medium, containing DAPI, was applied to each piece. In either case, the coverslip was sealed to the slide with nail polish and filters counted within 24 h.

Thirty ("OC" samples) and 50 ("WS" samples) 100 μm by 100 μm fields were enumerated per filter piece using a 100x oil-immersion objective on an Olympus BX61 epifluorescence microscope. Probe signal was detected in the FITC channel and associated DAPI fluorescence (showing the cell nucleus) verified during enumeration. The volume filtered, and area of the filter, were considered and cell concentrations calculated accordingly. Cells were placed into three size categories (using the largest cell dimension): <3 μm, 3-10 μm and >10 μm, by measurement against grid markings (1 μm

increments). More specific size measurements were performed as below. The number of cells in the >10 μm size fraction was statistically unreliable (0.8 ± 1.8%) and therefore not considered further.

Several controls were performed alongside PRYM02 hybridization of field samples. The bacterial antisense NON338 probe (5′ ACTCCTACGGGAGGCAGC 3′, Worden et al., 2000) was used as a negative control for all hybridizations. In addition, filters of *Micromonas sp.* RCC299 and *Isochrysis sp.* CCMP1244 cultures were used as negative and positive controls for the PRYM02 probe, respectively. These were hybridized alongside all field samples, including those from dilution experiments. A no-probe control was added for each environmental sample at least once, but not necessarily for each hybridization.

For all the "N" and "S" cruises, epifluorescence microscopy was used to enumerate and size prymnesiophytes cells. A known volume of surface water was added to a filter funnel fitted with a polycarbonate filter (Nuclepore, 25 mm diameter, 0.2 μm pore size) with a diffuser filter underneath, preserved with a small volume of 50% glutaraldehyde (1-2%, final concentration) and vacuumed onto the filter. Filter were mounted on glass slides in subdued light in order to preserve phytoplankton pigment fluorescence and counted aboard the ship on the day of collection using a Zeiss Axioskop equipped with epifluorescence and a 100x oil-immersion objective. The excitation filter was a Zeiss 48.77.09 under which phycoerythrin fluoresces orange and chlorophyll fluoresces red. Glutaraldehyde-induced green fluorescence revealed cell membranes and, in combination with pigment fluorescence, and the unique chloroplast and flagellar configurations of

prymnesiophytes, pico-prymnesiophytes were counted and sized. Pico-prymnesiophytes were binned to four size categories (Table 3.1). A 10 x 10 grid of 4,624 μm^2 was used to count abundant picophytoplankton. For lower abundances of picophytoplankton the iris diaphragm was closed to give a 120 μm diameter field and a portion of this field counted until >500 picophytoplankton had been routinely counted.

Picoplankton cell size and biomass

Cell size of prymnesiophytes was determined as above using epifluorescence microscopy. For these four pico-prymnesiophyte size categories (Table 3.1), biovolume was calculated using the formula $V=4/3\ \pi$ x (L/2) x (W/2) x (W/2). Because only two dimensions could be measured on the microscope, the third dimension for volume was assumed to be the shortest of the two dimensions measured (W), thus potentially biasing the data in a way that would underestimate pico-prymnesiophyte biovolume values. Biovolume for the pico-prymnesiophyte size categories ranged from 4.2 to 11.5 μm^3 (Table 3.1).

Biomass of various size groups was then estimated using the product of abundance and mean cellular carbon content. The latter was taken as the product of cell biovolume and a single carbon conversion factor used for all groups, 237 fg C μm^{-3}, previously reported for *Prochlorococcus*, *Synechococcus* and several 'non-prymnesiophyte' picoeukaryote groups (Worden et al., 2004). For pico-prymnesiophytes mean cellular carbon content was determined using this carbon conversion factor for each of the four size categories (Table 3.1). *Prochlorococcus* and *Synechococcus* cellular carbon

conversion values were 39 fg C $cell^{-1}$ and 82 fg C $cell^{-1}$, respectively, as determined previously on discrete populations enumerated by FCM and analyzed by CHN (Worden et al., 2004).

Table 3.1. Average size, biovolume and biomass conversion factor for various picoplankton groups. Cell measurements for four size classes of pico-prymnesiophytes binned during counting at all locations, except OC413 and Florida Straits. In the case of the Florida Straits, representing one biogeographical province data point (Figure 3.3), cells were binned into two size classes: <3 µm and >3 µm, the majority were <3 µm in their largest dimension. For the pico-prymnesiophytes classes, biovolume and then a biomass conversion factor was calculated from average size (numbers are rounded after calculation). More precise cell size information was available for the NSS and BATS (i.e., average within size class of 1.9 ± 0.4 x 2.1 ± 0.3 µm, n=89; 2.8 ± 0.6 x 3.4 ± 0.5 µm, n=161) resulting in slightly different biovolumes (4.0 and 14.0 μm^{-3}) than for the below class 1 and 4, respectively. The former were used to generate 3 of the 121 global data points, by averaging over the entire Florida Straits time series, and for Sargasso sites. Pico-prymnesiophyte values in all other tables or calculations refer to the sum of the individually calculated biomass for each group (i.e. biomass conversion factor x cell concentration). For *Prochlorococcus*, *Synechococcus*, 'non-prymnesiophyte' picoeukaryotes and the four pico-prymnesiophyte size classes, the same carbon factor per unit volume (237 fg C μm^{-3}, see Materials and Methods) was used (Worden et al. 2004).

Organism	Size (µm)	Biovolume (μm^3)	Biomass conversion factor (fg C $cell^{-1}$)
Prochlorococcus	*	*	39*
Synechococcus	*	*	82*
'non-prym' picoeukaryotes	*	*	530*
Pico-prym class 1	2.0 x 2.0	4.2	995
Pico-prym class 2	2.0 x 2.5	5.2	1,232
Pico-prym class 3	2.0 x 3.0	6.3	1,493
Pico-prym class 4	2.5 x 3.5	11.5	2,726

Pico-prym, pico-prymnesiophytes
* From Worden et al., 2004

As noted above, counts from our eukaryotic FCM analysis window for the smallest eukaryotes showed a tight correlation with the sum of all 'non-prymnesiophyte' red fluorescing picoeukaryotes counted by microcopy (Buck et al., 1996). Given that 'non-prymnesiophyte' picoeukaryotes (e.g. oceanic pelagophytes and prasinophytes, for instance *Ostreococcus* is ~1 μm diameter and *Micromonas* ~1.4-1.6 μm) tend to be smaller than the pico-prymnesiophytes, the biomass conversion factor 530 fg C $cell^{-1}$ was used for the FCM-enumerated 'non-prymnesiophyte' picoeukaryotes, again as determined previously for field populations containing many pico-prasinophytes in the eastern North Pacific (Worden et al., 2004). Unlike some studies which have used larger cellular conversion factors for eukaryotes that likely over estimate their contributions (Worden et al., 2004), the cellular carbon conversion factors used here for the four picophytoplankton groups (*Prochlorococcus*, *Synechococcus*, 'non-prymnesiophyte' picoeukaryotes and pico-prymnesiophytes) were derived using the same carbon per unit volume (237 fg C $cell^{-1}$, from Worden et al., 2004). This value is similar to that of Booth et al. (1988), 220 fg C μm^{-3}. However, Grob and colleagues (2007) reported *in situ* cellular carbon of *Prochlorococcus* being 29 ± 11 fg C $cell^{-1}$ and for combined picophytoeukaryotes (all lineages) being 730 ± 226 fg C $cell^{-1}$, based on a combination of culture based work, environmental Coulter Counter data and CHN measurements. The latter value likely reflects an average between our 'non-prymnesiophyte' picoeukaryote and pico-prymnesiophyte cellular carbon values. Our *Prochlorococcus* cellular carbon value is higher than estimated by Grob and colleagues (2007), but similar to that of Bertilsson et al. (2003). Total picophytoplankton carbon was taken as the sum of the

population biomasses for *Prochlorococcus*, *Synechococcus* and 'non-prymnesiophyte' picoeukaryotes and the various contributions of the different cells within the different pico-prymnesiophyte size ranges.

For a small portion of the data ("OC" and "WS" cruises), pico-prymnesiophytes were enumerated for three bins only: <3 μm, 3-10 μm and >10 μm. Because the mid-size category (3-10 μm) contained many cells in the smaller end of this range, we more precisely sized cells at two sites and two depths in the Sargasso Sea. PRYM02 FISH-hybridized cells were measured in two dimensions (L, length = the longest visible cell dimension; and W, width = the shortest visible cell dimension) using a calibrated sizing grid for the NSS station (15 m, n=60 and 70 m, n=60) and BATS station (15 m, n=60 and 75 m, n=60). To determine averages the data were placed into two size categories (with L <3 μm and those with L between 3 to <5 μm). Ten percent of the cells were ≥5 μm (largest dimension) and were not included for further analyses to avoid overestimation of pico-prymnesiophyte biomass. This resulted in an average cell length of 3.4 ± 0.5 μm (instead of 3.8 ± 1.1 μm, when including all cells >5 μm) and average width of 2.8 ± 0.6 μm (instead of 3.1 ± 1.0 μm, when including all cells >5 μm). The vast majority of pico-prymnesiophytes counted from "WS" cruises contained only <3 μm pico-prymnesiophytes (Supplementary Information Section II Figure S3.2 and S3.3). Pico-prymnesiophytes biovolumes for these samples ranged from 4.0 ± 1.0 μm^3 to 14.0 ± 3.6 μm^3 (Table 3.1).

High performance liquid chromatography

Samples for pigment analysis were obtained by filtering 1 l to 5 l of seawater depending on the depth and location through a 25 mm glass fiber filter (GF/F Whatman, Maidstone, UK). The filter was then placed in a cryovial and frozen in liquid nitrogen. For analysis, filters were thoroughly dried, placed in 3 ml of 90% acetone and vortexed for 45 s before placing them at -20 °C. After 24 h, filters were sonicated for 30 s and vortexed again for 45 s. The extract was then cleared through 0.8 μm filters. One ml of extract was mixed with 0.2 ml of 0.2 μm filtered autoclaved 18.2 MΩ water and placed in an Autosampler tray at 4 °C. HPLC hardware and analysis as previously described (Latasa et al., 2001). Chl *a*, as the sum of monovinyl (MVChl *a*) and divinyl (DVChl *a*) Chl *a*, was used as a measurement of total phytoplankton pigment biomass. *Prochlorococcus* contribution to Chl *a* was estimated directly as DVChl *a*. The contribution of the rest of major groups to MVChl *a* was quantified using Chemtax (Mackey et al., 1996), although a newer version that was provided to us prior to publication (Wright et al., 2009). Samples were initially separated in two subgroups: DCM and surface samples. The pigment dataset was carefully checked to distinguish the potential presence of a total of seven phytoplankton groups that could contribute to MVChl *a*: Prymnesiophyceae, Pelagophyceae, Prasinophytae, *Synechococcus*, Cryptophyceae, Dinophyceae and diatoms; with the following pigments: Chl c_2, peridinin, 19'-butanoyloxyfucoxanthin, fucoxanthin, prasinoxanthin, violaxanthin, 19'-hexanoyloxyfucoxanthin, alloxanthin and zeaxanthin. Among the distinguished groups zeaxanthin, the pigment marker of *Synechococcus*, also occurs in Prasinophytae and *Prochlorococcus*. According to the abundances of their pigment marker prasinoxanthin,

Prasinophytae were a minor group comparing to *Synechococcus* and *Prochlorococcus*. This result, together with the low concentration of zeaxanthin in Prasinophytae (Latasa et al., 2004), made the contribution of Prasinophytae to the zeaxanthin pool practically negligible. Therefore, it was considered that only *Synechococcus* and *Prochlorococcus* contributed significantly to the zeaxanthin pool. Because only the former group contributes to MVChl *a*, it is necessary to distinguish between $Zeax_{Syn}$ and $Zeax_{Pro}$ to apply Chemtax. We partitioned Zeax as $Zeax_{FCM} = Zeax\ Syn_{cell}^{-1} * [Syn]_{FCM} + Zeax\ Pro_{cell}^{-1} * [Pro]_{FCM}$, where Zeax Syn_{cell}^{-1} and Zeax Pro_{cell}^{-1} are the Zeax content per cell of *Synechococcus* and *Prochlorococcus*, respectively, and $[Syn]_{FCM}$ and $[Pro]_{FCM}$ were the *Synechococcus* and *Prochlorococcus* cell concentrations obtained from FCM for the same sample. Initial values for Zeax Syn_{cell}^{-1} and Zeax Pro_{cell}^{-1} were estimated by minimizing the $\sum (Zeax_{HPLC} - Zeax_{FCM})^2$ using the function Solver of Microsoft Excel in default mode (time = 100 s, Iterations = 100, Precision = 0.000001, Tolerance = 5, Convergence = 0.0001, Lineal estimation, Progressive derivative, Newton's method). We used a single, common Excel cell for all Zeax Syn_{cell}^{-1} with a seed value of 1.8 fg Zeax Syn_{cell}^{-1} as per (Kana and Glibert, 1987). The same procedure was applied for all Zeax Pro_{cell} samples but with a seed value of 1 fg Zeax Pro_{cell}^{-1} from (Cailliau et al., 1996). This procedure provides a single value of Zeax Syn_{cell}^{-1} and Zeax Pro_{cell}^{-1} for all the samples. A further refinement consisted of applying Solver a second time allowing the change of all the individual values of Zeax Syn_{cell}^{-1} and Zeax Pro_{cell}^{-1}. Prymnesiophytes have previously been categorized as falling into eight major pigment groups (Zapata et al., 2004). Some have pigment characteristics of diatoms (Type 1), others of diatoms with some additional minor pigments (Types 2, 3, 4, and 5), others which are a mixture of

more typical prymnesiophytes (Types 6 and 7) and one that has characteristics of Pelagophytes and Prymnesiophytes. Here we used Type 6 and 7 to represent prymnesiophytes, and Type 8 did not exist in the matrix, but rather was divided to Pelagophytes or to Prymnesiophyte Types 6 and 7.

Chemtax was applied following the procedures described in Latasa (2007) using the version 1.95 of Chemtax (Wright et al., 2009). Random pigment to Chl *a* ratios between 0.1 and 1 were used as seed values of 16 input matrices. Chemtax was run using the following parameters: ratio limits = 1000, initial step size = 50, step ratio = 2, epsilon limit = 0.0001, cutoff step = 30000, iterations limit = 1000, elements varied = 10 (number of pigments), subiterations = 1, Weighting = Bound relative (50). The output of each run was used as input for the following run and this procedure repeated 8 times. The median of each pigment ratio was incorporated to the final pigment ratio matrix. This matrix was then used to estimate the contribution of the different groups to MVChl *a* stock.

At sea growth rate experiments

Dilution experiments according to the methods of Landry and Hassett (1982) were conducted with modifications similar to those in Worden and Binder (2003). These experiments allowed us to estimate the growth and grazing mortality rates of natural phytoplankton populations (Landry and Hassett, 1982; Landry et al., 1995). Briefly, a series of bottles containing different ratios of raw seawater to filtered seawater were incubated for 24 h. For each experiment, triplicate bottles were prepared for the following dilution factors: 1.0, 0.6, 0.4, 0.3, 0.2, and 0.1 (the factors represent the fraction of raw seawater diluted with 0.2 μm filtered seawater for a final volume of 1 l). Bottles were

incubated in on-deck water baths for 24 h (from sunrise to sunrise). Water temperature was maintained using a flow-though system that constantly pumped surface seawater through the on-deck water baths. Therefore, surface (15 m) experiment was conducted at the *in situ* temperature and the 70 m experiment (DCM at 93 m) was incubated at a temperature <2 °C higher than at 70 m. *In situ* light and spectral conditions were simulated using a combination of blue and/or neutral density gel filters (Lee filters, Burbank, CA, USA) placed around the water bath.

FCM samples from each bottle were collected from the triplicate bottles at t = 0 h and t = 24 h, preserved and stored in liquid nitrogen, as were FISH samples. For FISH 180 ml was filtered, with replication, for each of the replicate 1.0 dilution treatment bottles, while 405 ml was filtered for the more dilute bottles. The net (apparent) growth rate in each bottle was calculated as the natural logarithm of the ratio of the final cell concentration to the initial concentration. Linear regression of the net growth rates against dilution factors was used to estimate the grazing mortality rates (g, slope) and growth rates (μ, y-intercept). Growth rates and grazing mortality rates of *Prochlorococcus*, *Synechococcus* and picophytoeukaryotes were estimated using FCM samples collected from each bottle. Prymnesiophytes were enumerated and sized using FISH on samples collected from 1, 0.4 and 0.2 dilution factors. We used an ANOVA to test the significance of the regression and report only rates from statistically significant data ($p = 0.06$, $r^2=0.73$, $p = 0.06$, $r^2=0.87$).

Phytoplankton primary production and grazing losses were calculated for the NSS station using the dilution experiment results. Abundance and biomass of picophytoplankton groups (determined by FCM or FISH as shown above) were used in

following calculations. Growth rates (μ, d^{-1}) and grazing mortality rates (g, d^{-1}) of *Prochlorococcus*, *Synechococcus*, 'non-prymnesiophyte' picoeukaryotes and pico-prymnesiophytes determined by dilution experiments (see above) were used to calculate primary production (i.e. PP, $\mu g\ C\ l^{-1}\ d^{-1}$) of each group using the following equations:

at any instant, t: $PPt = \mu * Bt$

where Bt ($\mu g\ C\ l^{-1}$) is the phytoplankton biomass at time t calculated with the following equation:

$$B_t = B_0 * [e^{(\mu-g)*t}]$$

so that PP integrated over the length of the experiment T (here 1 day), can be calculated with the following equations:

$$PP = \mu * B_0 * [e^{(\mu-g)*T} - 1] / (\mu-g*T)$$

where B_0 ($\mu g\ C\ l^{-1}$) is the initial biomass.

We also calculated biomass production without grazing mortality, which would represent the maximum potential PP (PP_{max}) of each group using the following equation:

$$PP_{max} = B_0 * [e^{\mu*T} - 1] / T$$

Results and Discussion

Development of a eukaryote-targeted metagenomics approach.

To facilitate future investigation of genomic features of uncultured pico-prymnesiophytes, we developed a targeted metagenomic approach for obtaining sequences directly from natural populations. Discrete photosynthetic picoeukaryote populations were discriminated by scatter and chlorophyll-derived autofluorescence characteristics and sorted from subtropical Gulf Stream waters using a high-speed cell sorter (Figure 3.1, Figure 3.2 inset). Whole genome amplification was performed on sorted populations by MDA method (Dean et al., 2002), similar to an approach recently used on a broadly distributed, uncultivated group of marine cyanobacteria (Zehr et al., 2008). An iterative approach involving 18S and 16S rRNA gene clone library sequencing from multiple samples using MDA-flow sort DNA, prior to the subtropical Gulf Stream sort advanced for metagenomic sequencing, allowed us to establish flow cytometric characteristics of pico-prymnesiophytes in oligotrophic waters. The targeted approach was necessary not only because most pico-prymnesiophyte taxa appear to be uncultivated, but also because marine microbial communities are highly diverse (DeLong and Karl, 2005; Sogin et al., 2006). This makes metagenomic methods that maintain connectivity of genetic material from an individual microbe particularly successful for identifying organism specific gene suites, and has led to discoveries of unexpected features in uncultured prokaryotes (DeLong and Karl, 2005; Zehr et al., 2008). Unicellular eukaryotes have larger genomes and lower gene density than bacteria and archaea, and are present at lower cell numbers in marine ecosystems, making efficient recovery of eukaryotic sequences difficult by standard metagenomic sampling

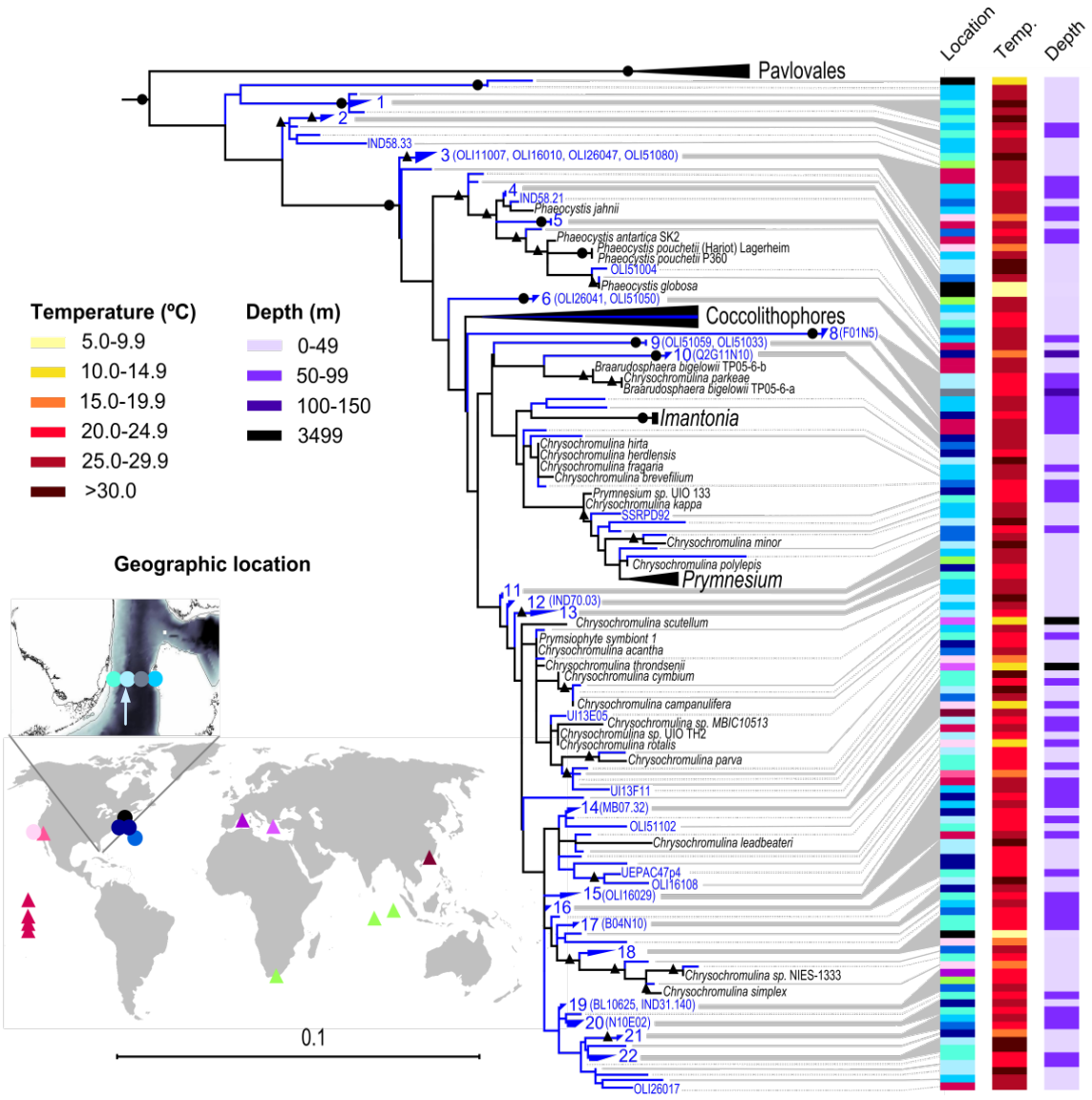
approaches such as seawater filtration. In addition, the paucity of relevant protistan genomes limits possibilities for accurate parsing of data from diverse communities. Thus, our targeted approach provided unprecedented access to genomic material from pico-prymnesiophytes by reducing organism diversity within the metagenome, and the bioinformatic complexity of data, at the outset.

Flow sorted pico-prymnesiophyte populations were evolutionarily distant from the nearest cultured taxa, but closely related to environmental sequences from native populations. PCR-based 18S rRNA gene clone libraries built from MDA-flow sort DNA from samples were analyzed within the broader context of <2 μm size-fractionated clone libraries generated from samples collected on multiple dates and cruises, that included the study region and nearby Sargasso Sea, and published data (Figure 3.2, Supplementary Information Section I Table S3.1). Similar to other studies (Moon-van der Staay et al., 2000; Fuller et al., 2006; Worden and Not, 2008; Liu et al., 2009) the vast majority of sequences were from uncultured prymnesiophytes (Figure 3.2, Supplementary Information Section I). In our 25 size-fractionated clone libraries, only a single sequence, seen on one date, was unambiguously assigned to a cultured species (100% identity to *Phaeocystis globosa*). 18S rDNA sequences from the Gulf Stream Station 04 MDA flow-sort (Figure 3.2, inset) showed two significant pico-prymnesiophyte clusters. 111 clones fell within environmental Group 8 at the 99% identity level, while 58 clones belonged to Group 15 (Figure 3.2). Environmental Group 3 was detected at low levels (7 clones). The largest cluster, Group 8, was also present in the Sargasso Sea (Not et al., 2007) as well as in the Florida Straits in summer and winter (Supplementary Information Section I Table S3.3). Overall, Group 8 remained unresolved in terms of phylogenetic placement (no

bootstrap support) and had only 93% SSU identity to any cultured organism (to members of the genus *Chrysochromulina*). This level of 18S rDNA divergence has important implications for differences in gene content. For example, the pennate and centric diatoms, *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, with 90% 18S rDNA identity, share only 30-40% of their total genes (Bowler et al., 2010). Of the four published picoeukaryote genomes (all prasinophytes, Palenik et al., 2007; Worden et al., 2009), the two sequenced *Micromonas* isolates, with 97% 18S rRNA gene identity, have 69.5% DNA identity over aligned genome regions, sharing at most 90% of their protein-encoding genes (Worden et al., 2009). Among cyanobacterial counterparts of picophytoeukaryotes, *Prochlorococcus* isolates have 98% 16S rRNA gene identity but extensive physiological differences based on genome and experimental analyses, occupying fundamentally different niches (Rocap et al., 2003).

A draft sequence of *E. huxleyi* is currently the only available prymnesiophyte genome. This larger species belongs to the coccolithophores, a unique group with ornate calcium carbonate plates. Based on 18S rDNA and other markers (Medlin et al., 2008), coccolithophores are distant from the soft-bodied prymnesiophytes within the flow-sort populations (Figure 3.2, Supplementary Information Section I). *E. huxleyi* and pico-prymnesiophytes are also expected to be unalike in terms of ecological niche, based on differences in cell size (Raven, 1998). Given the divergence between the flow-sorted population and cultured taxa, including *E. huxleyi*, application of a metagenomic approach to these samples will be extremely valuable for investigating genomic features of environmentally relevant pico-prymnesiophytes. This should allow identification of adaptations associated with the ecology of these organisms.

Figure 3.2 Maximum-likelihood reconstruction of 18S rDNA sequences from (blue lines) environmental samples and (black) cultured prymnesiophytes. Libraries from 25 discrete locations/depths sampled herein (circles, inset) were from the <2 µm size-fraction, except for the North Pacific (<3 µm size-fraction). Those from previous publications (triangles, inset) were primarily <3 µm size-fractionated. The Florida Straits (light blue circles) was sampled seasonally and the Gulf Stream Current core (blue arrow) determined using ADCP data. A single representative was used for redundant sequences within each library, resulting in 111 of our sequences being in the final phylogeny. Clades composed of environmental sequences with 99% nucleotide identity were collapsed after tree building (blue, Groups 1 to 22; Supplementary Information Section I). 18S rDNA sequences from green and red lineage organisms served as outgroups. Supported nodes are indicated for bootstrap percentages of (black circles) 100% and (black triangles) 75% or greater. Uncultured prymnesiophyte sequences have also recently been reported in South Pacific (Shi et al., 2009).



Abundance and growth in the study region

Despite their diversity and broad distribution, questions still remain regarding the ecological importance of pico-prymnesiophytes. Previous reports suggesting the importance of pico-prymnesiophytes in the open ocean require confirmation. The algorithms used for the pigment-analyses on which these reports are based, do not necessarily partition contributions by organism size, nor are samples collected for HPLC pigment analysis size fractionated. Furthermore, other lineages contain the prymnesiophyte-indicative marker pigment 19'-hexanoyloxyfucoxanthin (Carreto et al., 2001; Liu et al., 2009). Therefore, efforts paralleling the comparative genomic and evolutionary analyses of pico-prymnesiophytes focused on their importance to biomass and primary production, including specific growth rates.

We first investigated abundance and size distributions using microscopy with prymnesiophyte-specific FISH (Simon et al., 2000) on samples from the subtropical North Atlantic sites for which metagenome and clone library sequencing was performed (Figure 3.2). The majority of prymnesiophytes in the Florida Straits were picoplanktonic in size, in a transect sampled repeatedly over one year. 90 ± 9% and 87 ± 13% of prymnesiophytes were <3 μm in diameter in the surface and DCM, respectively (Supplementary Information Section II Figures S3.2 and S3.3). Results were similar in the Sargasso Sea, where more precise size measurements showed two primary size classes (Table 3.1). Direct count and cell size measurements were used to determine biomass contributions for the four major picophytoplankton groups: *Prochlorococcus*, *Synechococcus*, pico-prymnesiophytes and 'non- prymnesiophyte' picoeukaryotes (Buck et al., 1996, Table 3.2 and 3.3, Supplementary Information II Figures S3.2, S3.4, S3.5

and S3.6), with the latter often considered composed of prasinophytes and pelagophytes (Vaulot et al., 2008; Worden and Not, 2008). Although more nutrients were presumably available at the DCM, pico-prymnesiophyte abundance was typically higher in surface waters as was their percent contribution to picophytoplankton biomass, in the Florida Straits. Contributions in the Sargasso Sea were roughly equivalent between the surface and DCM (23 and 21% of picophytoplankton carbon, respectively; Supplementary Information Section II). The direct-count based biomass approach resulted in lower pico-prymnesiophyte biomass contributions at the DCM than estimated by HPLC (Supplementary Information II Figures S3.7-S3.9). However, values from the two methods were similar for surface waters, supporting the long-held HPLC-based inference that in open ocean surface waters the bulk of prymnesiophyte biomass is contributed by picoplanktonic taxa.

Pico-prymnesiophytes were capable of high growth rates, which could amplify relative contributions to primary production. Pico-prymnesiophyte specific growth rates were established for the first time using the dilution method and direct counts in the subtropical Atlantic (Landry, 1993). This method renders growth rates close to those measured by *in situ* cell cycle analysis for taxa amenable to the latter analysis, i.e. *Prochlorococcus* (Worden and Binder, 2003). Pico-prymnesiophyte growth rates were high (1.12 d^{-1}, r^2=0.87, $p < 0.07$) at the surface and lower (0.29 d^{-1}, r^2=0.73, $p < 0.07$) at 70 m, above the DCM depth (93 m) in the Sargasso Sea. Groups 13, 15 and 16 were detected (Figure 3.2). The *Prochlorococcus* growth rate was lower (0.63 d^{-1}, r^2=0.54, $p < 0.01$) than that of the pico-prymnesiophytes at the surface, but higher (0.60 d^{-1}, r^2=0.61, $p < 0.001$) at depth. Because this constitutes the first report of pico-prymnesiophyte specific growth

measurements, *Prochlorococcus* rates were used to ascertain whether the rates herein were comparable to literature reports. Rates were similar to previous direct count-based *Prochlorococcus* growth rates in the region (0.52 d^{-1}) for the same time of year (Worden and Binder, 2003), but higher than estimates obtained by radioactive label incorporation (~0.4 d^{-1}, Goericke and Welschmeyer, 1998). Furthermore, unadjusted total (all size fractions) prymnesiophyte HPLC ratio-based dilution growth estimates in the equatorial Pacific, a region dominated by picophytoplankton (Landry et al., 2003), are similar to those herein. Combined growth rate, mortality and biomass data were used to estimate primary production. Despite lower relative abundance, the combination of greater cellular biomass and a high growth rate led to pico-prymnesiophyte primary production levels comparable to that of *Prochlorococcus* in surface waters (1.1 and 1.8 μg C l^{-1} d^{-1}, respectively).

Latitudinal gradients and global contributions

Together, these data indicate that pico-prymnesiophytes contribute a major portion of primary production in the subtropical North Atlantic. Still, variations between ocean basins, as well as latitudinal gradients, translate to major biotic differences (Van Mooy et al., 2009) and the respective microbial communities will likely experience, and respond to, environmental change differently. Global surface biomass contributions by pico-prymnesiophytes were investigated using direct count and cell size data from 11 oceanographic expeditions, including those in the Sargasso Sea and Florida Straits (Table

3.1). Five biogeographical provinces were evaluated, the tropics (low latitudes) as well as subtropical-temperate (mid-latitudes) and sub-polar (high-latitudes) systems for both hemispheres (Table 3.2 and 3.3).

Pico-prymnesiophytes formed a substantial portion of the total picophytoplankton carbon globally and their contributions showed a strong latitudinal signal (Figure 3.3, Table 3.2). In high latitudes, pico-prymnesiophytes dominated all picophytoplankton, comprising 50-56% of the biomass in sub-polar provinces. Pico-prymnesiophyte biomass ranged from about 0.5-2.3 fold that of 'non-prymnesiophyte' picoeukaryotes in these cold waters and *Prochlorococcus* was effectively absent (Table 3.2 and 3.3). In mid-latitude provinces, pico-prymnesiophyte biomass levels were equivalent, but variations in other picophytoplankton groups modified the relative proportion of their contributions. Amongst picoplanktonic eukaryotes they comprised 66% and 33% in the northern and southern subtropical-temperate provinces, respectively, possibly reflecting seasonal or other environmental influences on community composition. The large pico-prymnesiophyte biomass contributions again resulted in part from their cell size and consequent biovolume (Table 3.1), which was greater than other picophytoplankton taxa, particularly cyanobacteria and prasinophytes (Worden and Not, 2008). Thus, the significance of their contributions is less obvious when abundance alone is considered (Table 3.3); *Prochlorococcus*, for instance, is orders of magnitude more abundant than the pico-prymnesiophytes in low- and mid-latitudes, but is also significantly smaller.

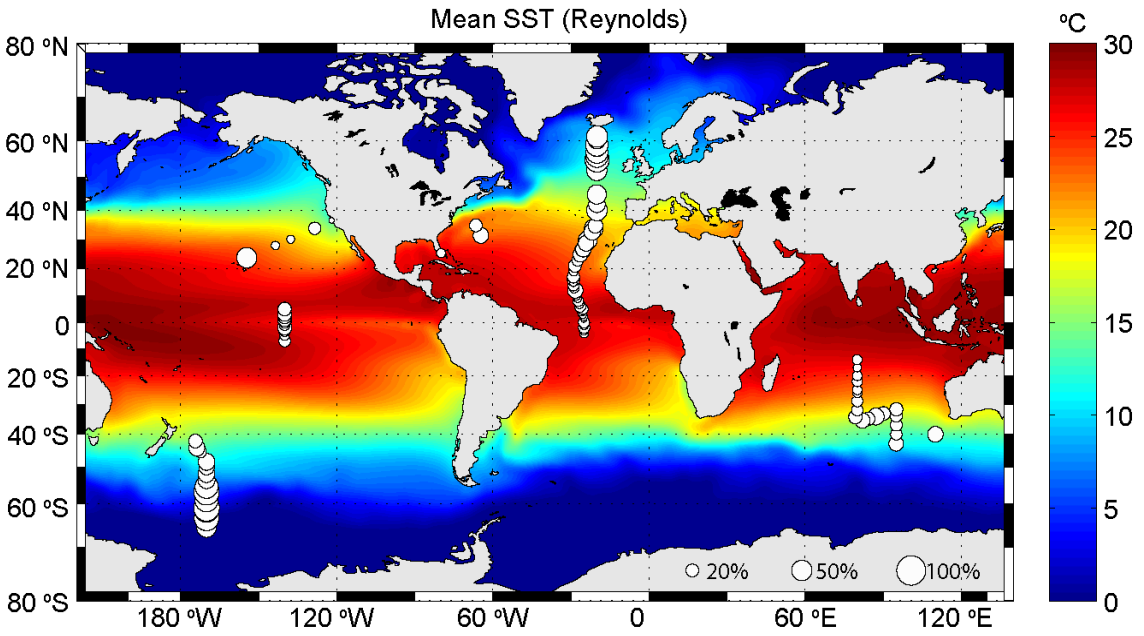
Table 3.2. Average biomass and biomass range for each picophytoplankton group. Surface sample average and standard deviation (adjacent parentheses) as well as range of cells ml^{-1} for each latitudinal zones is shown for *Prochlorococcus*, *Synechococcus*, 'non-prymnesiophyte' picoeukaryotes, pico-prymnesiophytes and total picophytoplankton (i.e. sum of the four groups).

	Latitude	**Average ($\mu g\ C\ l^{-1}$)**	**Range ($\mu g\ C\ l^{-1}$)**
Prochlorococcus	65º N-45º N	0.1 (0.1)	0.0-0.4
	45º N-20º N	3.7 (3.0)	0.1-12.8
	20º N-20º S	8.4 (2.8)	1.7-20.3
	20º S-45º S	3.1 (2.0)	0.5-9.5
	45º S-65º S	0.3 (0.4)	0.0-1.4
Synechococcus	65º N-45º N	2.4 (1.1)	1.4-4.6
	45º N-20º N	0.5 (0.5)	1.0-2.6
	20º N-20º S	1.0 (1.1)	0.0-6.6
	20º S-45º S	0.4 (0.6)	0.0-2.8
	45º S-65º S	1.1 (1.9)	0.00-5.5
'non-prym' picoeukaryotes	65º N-45º N	3.0 (1.7)	0.7-5.5
	45º N-20º N	1.0 (2.7)	0.01-13.2
	20º N-20º S	1.8 (1.7)	0.0-6.7
	20º S-45º S	3.7 (1.4)	0.9-6.0
	45º S-65º S	3.9 (2.8)	0.8-8.1
pico-prymnesiophytes	65º N-45º N	7.0 (4.7)	1.8-16.4
	45º N-20º N	2.0 (1.5)	0.4-5.3
	20º N-20º S	1.8 (0.9)	0.7-4.3
	20º S-45º S	1.8 (0.9)	0.6-3.4
	45º S-65º S	5.5 (4.9)	1.6-16.6
Total	65º N-45º N	12.5 (5.9)	5.7-23.7
	45º N-20º N	7.2 (4.2)	2.3-19.5
	20º N-20º S	13.0 (3.9)	7.4-29.7
	20º S-45º S	9.0 (3.0)	4.5-14.8
	45º S-65º S	10.9 (7.2)	3.2-21.7

Table 3.3. Average abundance and abundance range for each picophytoplankton group. Surface sample average and standard deviation (adjacent parentheses) as well as range of cells ml^{-1} for each of the biogeographical provinces is shown for *Prochlorococcus*, *Synechococcus*, 'non-prymnesiophyte' picoeukaryotes, pico-prymnesiophytes and total picophytoplankton (i.e. sum of the four groups).

	Latitude	**Average (cells ml^{-1})**	**Range (cells ml^{-1})**	**Number of samples**
Prochlorococcus	65º N-45º N	3,063 (2,902)	278-9,286	8
	45º N-20º N	94,176 (77,794)	2,434-328,550	24
	20º N-20º S	215,798 (70,706)	43,828-519,792	59
	20º S-45º S	78,984 (52,438)	11,659-243,197	19
	45º S-65º S	8,477 (11,331)	517-36,817	11
Synechococcus	65º N-45º N	29,189 (13,530)	17,410-56,299	8
	45º N-20º N	6,173 (6,624)	983-31,394	24
	20º N-20º S	12,682 (13,188)	442-79,899	59
	20º S-45º S	5,168 (7,903)	241-33,732	19
	45º S-65º S	13,694 (23,197)	0-66,648	11
'non-prym' picoeukaryotes	65º N-45º N	5,664 (3,159)	1,392-10,449	8
	45º N-20º N	1,937 (5,039)	9 -24,925	24
	20º N-20º S	3,353 (3,233)	29-12,665	59
	20º S-45º S	6,988 (2,712)	1,645-11,386	19
	45º S-65º S	7,434 (5,300)	1,567-15,240	11
pico-prymnesiophytes	65º N-45º N	2,933 (2,077)	757-7,144	8
	45º N-20º N	815 (482)	233-1,801	24
	20º N-20º S	784 (360)	378-2,063	59
	20º S-45º S	936 (474)	348-2,052	19
	45º S-65º S	2,017 (1,801)	590-6,122	11
Total	65º N-45º N	40,849 (13,035)	24,685-59,756	8
	45º N-20º N	104,614 (74,694)	18,684-335,715	24
	20º N-20º S	232,618 (76,831)	62,345-603,579	59
	20º S-45º S	92,076 (51,163)	19,133-248,687	19
	45º S-65º S	31,622 (37,596)	19,133-119,709	11

Figure 3.3. Global surface biomass contributions of prymnesiophytes as a percentage of total picophytoplankton carbon. Percent of pico-prymnesiophyte contributions to the total is represented by bubble size (scaling in the lower right of panel). Sea Surface Temperature (SST) represents 1^{o} increments averaged monthly over 18 years, highlighting differences in the five biogeographical provinces.



Our *in situ* observations were at odds with a recent report suggesting pico-prymnesiophytes are of greater importance than *Prochlorococcus* in the tropics (Liu et al., 2009). Were the latter report confirmed, it would result in major reconsideration of primary production and community dynamics in the tropics. However, our direct count-based data showed a different distribution than estimated in that study (Liu et al., 2009), for which an algorithm was used that relates satellite surface chlorophyll to prymnesiophyte pigments and their contribution to total chlorophyll. We found that pico-prymnesiophytes contributed 1.8 µg C l^{-1}, which corresponded to one fifth (21%) of the

biomass contributed by *Prochlorococcus* in low latitudes (Table 3.2). This significant discrepancy may result from issues surrounding the algorithm-based approach, such as: 1) the potential contribution of dinoflagellates, which can be abundant in the tropics (Landry et al., 2003) and can contain prymnesiophyte indicative pigments (Carreto et al., 2001), 2) a variable relation between a specific pigment content and surface chlorophyll, or 3) the fact that pigment-based algorithms do not partition contributions by organism size. Our results, based on direct-enumeration of each of the major picophytoplankton groups, rendered results for *Prochlorococcus* that corresponded well with extensive HPLC data on *Prochlorococcus* surface biomass contributions in the tropics and subtropics (Letelier et al., 1993; Landry et al., 2003).

The data show pico-prymnesiophytes are highly successful in oligotrophic settings. However, it is not surprising that *Prochlorococcus* dominates in tropical systems which are typically highly oligotrophic. *Prochlorococcus* presumably has a considerable competitive advantage in systems that are stratified for extended periods, like tropical oceans, given its genome streamlining, much smaller cell size and low cellular quotas for some limiting nutrients (Rocap et al., 2003). Features within the targeted pico-prymnesiophyte metagenome still remain difficult to relate to niche differentiation, given the plethora of predicted genes of unknown function (37%), similar to results for other genomes and metagenomes. This lack of functional understanding is perhaps the greatest impediment to connecting genomes to organism physiology and response or competition processes. Our data positions us to now explore the function of such genes *in situ*, through experimentation and additional field measurements of these uncultured phytoplankton.

Globally, pico-prymnesiophytes averaged 2.6 ± 1.8 μg C l^{-1} when the areal extent of each province (Table 3.4) was accounted for, and, although less evenly distributed given its near absence from cold waters, *Prochlorococcus* averaged 4.7 ± 2.1 μg C l^{-1}. Pico-prymnesiophyte biomass per liter was maximal in the northern hemisphere sub-polar province, but the massive extent of the Southern Ocean, relatively unimpeded by land, rendered their greatest contributions to global biomass in the southern hemisphere sub-polar province. Thus, tiny prymnesiophytes composed a significant proportion of primary production, at least in the Sargasso Sea, and global picophytoplankton biomass.

Table 3.4. Average surface biomass of the four picophytoplankton groups per liter of water for five biogeographical provinces from 121 values, some composed of multiple data points from seasonal sampling and all by direct count with cell sizing of eukaryotes in two dimensions. Area varies significantly over the latitudinal zones due to the influence of different land masses. Standard deviations are provided in parentheses adjacent to the relevant value.

Latitude	Ocean area (x 10^{12} m^2)	Samples (number)	Biomass (μg C l^{-1})			
			Prochlorococcus	*Synechococcus*	"non-prym" picoeukaryotes	pico-prymnesiophytes
60º N-45º N	13.22	8	0.1 (0.1)	2.4 (1.1)	3.0 (1.7)	7.0 (4.7)
45º N-20º N	47.18	24	3.7 (3.0)	0.5 (0.5)	1.0 (2.7)	2.0 (1.5)
20º N-20º S	122.60	59	8.4 (2.8)	1.0 (1.1)	1.8 (1.7)	1.8 (0.9)
20º S-45º S	75.35	19	3.1 (2.0)	0.4 (0.6)	3.7 (1.4)	1.8 (0.9)
45º S-65º S	48.69	11	0.3 (0.4)	1.1 (1.9)	3.9 (2.8)	5.5 (4.9)

Conclusions

Soft bodied prymnesiophytes survived the mass extinctions of the K/T (Cretaceous-Tertiary) boundary (Medlin et al., 2008), indicating that, in past extinction events, prymnesiophyte taxa akin to those analyzed herein were resilient to perturbation.

Increased stratification, reduced mixing and lower nutrient concentrations predicted under some climate-change scenarios for present-day oceans, are hypothesized to create conditions favoring picoplankton (Worden et al., 2009). Based on the limited data available, such population shifts were recently reported in the Arctic Ocean (Li et al., 2009), and presumably will impact sub-polar and warmer water systems as well. The success of tiny uncultured prymnesiophytes in modern oceans, and their contributions to primary production in future times, are linked to evolutionary history and genetic make-up, as well as the rate and extent of perturbations experienced.

The data herein, showing the significant global contributions of pico-prymnesiophytes provides several advancements. The diversity of marine microbial communities makes it difficult to determine which populations are most critical for input in global biogeochemical models and research initiatives focused on ecosystem level processes. In turn, this also shapes priorities for *in situ* sensor development, an essential step for improving spatial and temporal resolution for marine studies; our work highlights the need to prioritize pico-prymnesiophytes.

Supplementary Information Section I

This supporting information contains more detailed information on small subunit phylogenetics. Results are discussed to a greater extent within the context of previous literature. Also provided are tables providing sample locations and dates for 18S rRNA clone libraries (Supplementary Information Section I Table S3.1), GenBank accession numbers for published sequences used in Figure 3.2 (Supplementary Information Section I Table S3.2) and composition of environmental sequence Groups shown in (Figure 3.2 Supplementary Information Section I Table S3.3). The 18S rRNA gene tree as in Figure 3.2 is also provided showing clone names (Supplementary Information Section I Figure S3.1).

The overall 18S rRNA gene tree used new sequences from 26 clone libraries (i.e. 25 discrete locations/depths, all being from <2 μm or for one library <3 μm pore-size pre-filtered seawater, one additional sequence was included that was derived from cells collected on a 2 μm pore size filter; Supplementary Information Section I Table S3.1), flow sorted samples from two sites, as well as previously published clone libraries and representatives of all cultured prymnesiophytes sequenced as of February 2009 of suitable quality (retrieved from GenBank, Supplementary Information Section I Table S3.2). Representative sequences from the flow-sort MDA clone libraries, including two additional flow sorts (that were not advanced for metagenomic sequencing) were included (Figure 3.2, Supplementary Information Section I Figure S3.1). In order to simplify the presentation of the overall 18S rRNA gene tree, sequences within individual clades that had 99% identity to each other were collapsed, after phylogenetic

reconstruction, and assigned a group number. This also provided a relatively easy way to compare our sequences to those from previous studies.

The topology of the overall 18S rRNA tree (Figure 3.2, Supplementary Information Section I Figure S3.1) was consistent with previous reports. In general, bootstrap values at interior nodes were low as commonly seen for 18S rRNA gene trees. Node support was especially low for the Prymnesiales, and most deep branches in our phylogenetic reconstruction (Figure 3.2) were unresolved. SSU rRNA gene phylogenies are known to have limited resolution and hence while the trees clearly demonstrate the extensive diversity of uncultured taxa within the prymnesiophytes, evolutionary relationships are difficult to discern. The Pavlovales formed a supported clade distinct from, and basal to, the prymnesiophytes and several of our clone library sequences were placed basal to cultured prymnesiophytes, but inside the Pavlovales, as seen elsewhere (Takano et al., 2006). Several major clades within the prymnesiophytes were represented, which in the past have been placed into extremely broad "Clades", specifically Clades A to E (Edvardsen et al., 2000; Sáez et al., 2004; Edvardsen and Medlin, 2007). Clade A contained all *Phaeocystis*, a broadly distributed genus. We retrieved only one related sequence, from the NSS (150 m), which had 100% identity to *P. globosa*, a common species in temperate and subtropical waters and cell size around 3 to 4.5 μm (Medlin and Zingone, 2007). *P. globosa* formed a clade with *P. pouchetii* (Arctic) and *P. antarctica* (Antarctica, Lange et al., 2002), both polar species. Our environmental sequences were not closely related to either *P. pouchetii* or *P. antarctica*, in agreement with their geographical distribution. Most of the environmental sequences retrieved from our survey that were placed within Clade A fell outside the *P. globosa* clade, as did the

Mediterranean *P. jahnii*, a species found in warm waters. The non-colony forming *P. cordata*, also isolated from the Mediterranean Sea, was not included in our analysis due to ambiguities in the available deposited sequence. However, none of our environmental sequences had more than 97% identity to this species.

More than half (58%) of our environmental sequences fell within Clade B and, specifically, Clade B2 (B1 and B2 are broad subgroups within B). Clade B contains the non-mineralized order Prymnesiales including *Chrysochromulina*, *Prymnesium* and *Imantonia*. Nevertheless, these environmental sequences did not have close cultured representatives and formed a series of sub-groups themselves. The fact that the environmental sequences did not clade with cultured strains is particularly interesting in view of the fact that about 50 species having been described within the genus *Chrysochromulina*, most of which are marine (Jordan et al., 2004), and many of which have deposited 18S rRNA gene sequences. Clade B2 included sequences from all regions sampled (worldwide locations). A recent study has shown that B2 is present in high latitudes based on clone library analyses (Liu et al., 2009). Furthermore, uncultured pico-prymnesiophyte sequences have also recently been reported in the South Pacific (Shi et al., 2009). In addition, analysis of 16S genes has also shown many uncultured taxa (Rappé et al., 1998; Fuller et al., 2006).

Only five of the environmental sequences we retrieved fell within the coccolithophores (Clade C) and none were closely affiliated with sequenced taxa. Several of these were placed in Group 7 (FS04E051_31Mar05_5m and FS14K029_31July05_5m, clustered at 99% identity), while others were singletons (FS14JA16_30Mar05_5m, FS14JA75_30Mar05_5m, FS14M081_08Dec05_58m). That so few Clade C

representatives were retrieved was not surprising for two reasons. First our size fractionation would generally exclude many of these cells. Secondly, our cell size measurements of the natural populations showed there were very few cells within an appropriate size range for this Clade. For example, in the Florida Straits, of the 36 FISH samples analyzed only eight had >15% of the prymnesiophyte cells falling within the >3 μm size fraction, the majority were <3 μm in size. In these eight samples, 28 ± 10% fell were in the 3-10 μm size fraction, with the rest being smaller (Supplementary Information Section II Figures S3.2 and S3.3).

Group 3 and Group 6 include a number of subtropical Atlantic sequences as well as a number of previously published sequences from the equatorial Pacific. Those from the previous study (Moon-van der Staay et al., 2000) were OLI16010, OLI11007, OLI51080 and OLI26047, which have been termed Clade D in (Edvardsen et al., 2000). Group 6 (Figure 3.2) included sequences from a formerly defined Clade E (OLI26041 and OLI51050, Edvardsen et al., 2000). Group 8 had longer branch lengths than most identified clades and did not retain bootstrap support at inner nodes. Furthermore, its overall placement was quite unstable in preliminary analyses. Sequences composing Group 8 had previously been reported in the Sargasso Sea (Not et al., 2007), which were then recovered herein again from the Sargasso Sea, the seasonal Florida Straits samples, and the flow sort-MDA clone libraries. We also explored seasonal aspects of diversity by sampling over the course of the year (6 cruises evaluated by FCM and FISH, 5 with HPLC and 3 for which clone libraries were built) in the Florida Straits. We were not able to discern a seasonal trend in the pico-prymnesiophyte sequences retrieved, although the statistical depth of sampling was low, which could have obscured patterns.

Finally, a recent study exploring label uptake in eukaryotes from labeled *Prochlorococcu*s prey (presumably uptake was direct) showed that sequences close to Group 14 (having 99% nucleotide identity to FS04R13_10_75m_sort, Supplementary Information Section I Figure S3.1, Supplementary Information Section I Table S3.3) were present at Station ALOHA, in the north Pacific Gyre, e.g. hotxp4g5 (Frias-Lopez et al., 2009). Thus, a consideration regarding some uncultured prymnesiophytes lies in emerging evidence that some may be capable of consuming *Prochlorococcus* (Frias-Lopez et al., 2009). Several sequences recovered in the study by Frias-Lopez et al. (2009) bore 99% identity to a sequence close to Group 14 (Figure 3.2), however not to the environmental groups represented in the flow sort. In our study, the pico-prymnesiophytes evaluated contained chlorophyll and showed no evidence of captured prey. Some potential prey, like *Synechococcus*, would be difficult to overlook. However, if other pico-prymnesiophyte groups can switch trophic modes facultatively, it would dictate whether they serve as primary producers versus surviving by energy capture through predation, or some mixture of the two modes. The environmental triggers behind such a switch would reshape food web dynamics and their ecosystem roles – as well as the capacity to deduce function from SSU sequences.

A note on terminology: the term prymnesiophytes is used to refer to the class Prymnesiophyceae Hibberd, this seems to be the most consistent usage in comparison to previous oceanographic literature. In general, this group is alternatively referred to as the division Haptophyta (division Haptophyta Hibberd ex Edvardsen et Eikrem), including both the Prymnesiophyceae and the Pavlovophyceae (Cavalier-Smith) Green et Medlin. It

should be noted however that classification of haptophytes has differed noticeably between authors (e.g. Parke and Dixon, 1976; Chrétiennot-Dinet et al., 1993; Green and Jordan, 1994, Jordan et al., 2004). Overall we adopted the nomenclature of Edvardsen et al. (2000) wherein: coccolithophores, incorporates all haptophytes with calcified scales (coccolithophores) during some stages of their life cycle (Fujiwara et al., 2001), includes two orders: Coccolithales (E. Schwarz) Edvardsen et Eikrem and Isochrysidales (Pascher) Edvardsen et Eikrem. The order Isochrysidales further divides into two families: Noëlaerhabdaceae Jerkovic including *Emiliania huxleyi* and *Gephyrocapsa oceanica*, and Isochrysidaceae (Bourrelly) Edvardsen et Eikrem including organisms that do not have mineralized scales (presumably having lost the ability to calcify) such as *Isochrysis galbana*. The order Coccolithales is also formed of a diverse array of families.

Table S3.1. Coordinates, dates and environmental characteristics of 18S rRNA gene clone library sites.

Location	Lat (N)	Lon (W)	Date (d/m/y)	Sample depth (m)	Temp (ºC)	Sal (ppt)
BATS	31º39'20"	64º37'21"	29/05/05	75	20.08*	36.79*
BATS	31º39'20"	64º37'21"	01/06/05	15	25.51*	36.68*
NSS	35º09'24"	66º33'46"	05/06/05	15	21.4-21.6*	36.5-36.6*
NSS	35º09'24"	66º33'46"	08/06/05	70	20.1	36.6
NSS	33º14'07"	64º53'19"	07/04/01	150	18.6	36.6
NSS	36º07'30"	67º10'03"	08/04/01	4	21.9	36.5
NSS	33º14'07"	64º53'19"	07/04/01	15	18.7	36.6
CS	40º15'07"	70º25'23"	09/04/01	4	5.6	32.7
CS	40º15'07"	70º25'23"	09/04/01	20	5.3	32.7
CS	39º59'38"	71º48'01"	16/09/01	26	14.4	33.3
FS St01	25º30'07"	80º04'04"	30/03/05	5	24.3	36.4
FS St01	25º30'07"	80º04'04"	30/03/05	70	21.2	36.4
FS St01	25º30'04"	80º03'59"	01/08/05	5	30.1	36.1
FS St01	25º30'04"	80º03'59"	01/08/05	65	23.5	36.3
FS St04	25º30'01"	79º57'20"	31/03/05	5	24.7	36.3
FS St04	25º30'04"	79º57'18"	01/08/05	5	30.3	36.0
FS St04	25º30'04"	79º57'18"	01/08/05	89	24.2	36.5
FS St04**	25º30'04"	79º57'18"	27/02/07	75	23.3	36.7
FS St08**	25º18'00"	79º34'12"	27/02/07	141	23.25	36.7
FS St14	25º29'59"	79º20'58"	30/03/05	5	25.7	36.2
FS St14	25º29'59"	79º20'58"	30/03/05	70	24.7	36.6
FS St14	25º29'55"	79º20'54"	31/07/05	5	29.9	36.0
FS St14	25º29'55"	79º20'54"	31/07/05	80	25.9	36.5
FS St14[§]	25º30'01"	79º21'04"	08/12/05	58	26.2	36.3
NE Pac[¶]	36º07'34"	123º29'24"	01/10/07	10	16.1	33.0
NE Pac	33º17'12"	129º25'41"	07/10/07	10	19.0	33.2
NE Pac[¶]	33º17'12"	129º25'41"	07/10/07	90	13.7	33.1

*Parameters measured with CTD detector prior to the sample collection (if single measurement) or prior to and the following day (if range of measurements), as samples were collected with GO-FLO bottle not equipped with a CTD detector.
BATS, Bermuda Atlantic Time-series Study; NSS, Northern Sargasso Sea; FS, Florida Straits; St, Station; CS, Continental Shelf, NE Pac, North East Pacific Ocean.
[§]2 clone libraries, 2 size fractions: < 2 μm and > 2 μm.
[¶]2 clone libraries.
**From MDA-flow sort DNA

Table S3.2. Taxonomic affiliation or environmental clone identifier and accession number of 18S rRNA gene sequences used in Figure 3.2. When available, strain information is provided.

Species/environmental clone	Strain or isolate	Accession Number	Supergroup	First Rank
Chondrus crispus		Z14140	Archaeplastida	Rhodophyceae
Gracilaria lemaneiformis		M54986	Archaeplastida	Rhodophyceae
Compsopogon coeruleus	SAG B 36.94	AF342748	Archaeplastida	Rhodophyceae
Cryptomonas ovata		EF180057	Chromalveolata	Cryptophyceae
Cryptomonas pyrenoidifera	CCAP 979/61	AJ421147	Chromalveolata	Cryptophyceae
Hemiselmis virescens	CCMP 443	AJ007284	Chromalveolata	Cryptophyceae
Rhodomonas salina	CCAP 978/24	EU926158	Chromalveolata	Cryptophyceae
Pyrenomonas helgolandii	SAG 28.87	AB240964	Chromalveolata	Cryptophyceae
Pyramimonas australis		AJ404886	Archaeplastida	Chloroplastida
Chlamydomonas reinhardtii	CC-1418	AY665726	Archaeplastida	Chloroplastida
Chlorella vulgaris	CCAP 211/11F	AY591515	Archaeplastida	Chloroplastida
Micromonas	CCMP1723	AY954997	Archaeplastida	Chloroplastida
Micromonas pusilla	CCMP1545	AY954994	Archaeplastida	Chloroplastida
Symbiodinium microadriaticum	NEPCC737	EF492496	Chromalveolata	Alveolata
Prorocentrum micans		AJ415519	Chromalveolata	Alveolata
Karlodinium micrum	NEPCC734	EF492506	Chromalveolata	Alveolata
Coscinodiscus radiatus	CCMP 309	X77705	Chromalveolata	Stramenopile
Thalassiosira weissflogii	CCMP1049	AY485445	Chromalveolata	Stramenopile
Thalassiosira pseudonana	CCMP 1007	DQ093367	Chromalveolata	Stramenopile
Gloeochaete wittrockiana	SAG B 46.84	X81901	Archaeplastida	Glaucophyta
Glaucocystis nostochinearum	SAG 45.88	X70803	Archaeplastida	Glaucophyta
Cyanophora paradoxa	UTEX 555	AY823716	Archaeplastida	Glaucophyta
Corcontochrysis noctivaga	AC 88	DQ207406	Chromalveolata	Haptophyta
Diacronema vlkianum	CCMP 504	AF106056	Chromalveolata	Haptophyta
Pavlova gyrans	CCMP 607	U40922	Chromalveolata	Haptophyta
Pavlova sp.	MBIC10094	AB183588	Chromalveolata	Haptophyta
Pavlova sp.	MBIC10665	AB183627	Chromalveolata	Haptophyta
Pavlova pinguis		AB293551	Chromalveolata	Haptophyta
Pavlova sp.	MBIC10455	AB183598	Chromalveolata	Haptophyta
Pavlova pinguis	MBIC10458	AB183600	Chromalveolata	Haptophyta
Pavlova pinguis	IIY089	AF106058	Chromalveolata	Haptophyta
Phaeocystis jahnii		AF163148	Chromalveolata	Haptophyta
Phaeocystis antarctica Karsten	SK23	X77481	Chromalveolata	Haptophyta
Phaeocystis pouchetii		X77475	Chromalveolata	Haptophyta
Phaeocystis pouchetii	P360	AF182114	Chromalveolata	Haptophyta
Phaeocystis globosa		EU127475	Chromalveolata	Haptophyta
Scyphosphaera apsteinii		AM490984	Chromalveolata	Haptophyta
Algirosphaera robusta	ALGO Am 24	AM490985	Chromalveolata	Haptophyta
Helicosphaera carteri	ALGO NS1010	AM490983	Chromalveolata	Haptophyta
Coronosphaera mediterranea	ALGO NS85	AM490986	Chromalveolata	Haptophyta
Chrysoculter rhomboideus		AB158370	Chromalveolata	Haptophyta
Syracosphaera pulchra	ALGO GK 17	AM490987	Chromalveolata	Haptophyta
Gephyrocapsa oceanica	PLY G01	AJ246276	Chromalveolata	Haptophyta
Gephyrocapsa oceanica	MBIC10537	AB058360	Chromalveolata	Haptophyta
Gephyrocapsa oceanica	MBIC11100	AB183665	Chromalveolata	Haptophyta
Emiliania huxleyi		L04957	Chromalveolata	Haptophyta
Holococcolithophorid sp.	ALGO holo	AM490989	Chromalveolata	Haptophyta
Helladosphaera sp.	ALGO Niesh	AM490991	Chromalveolata	Haptophyta
Calyptrosphaera radiata	ALGO P80-5	AM491024	Chromalveolata	Haptophyta
Calyptrosphaera sp.	ALGO Calyp2	AM490988	Chromalveolata	Haptophyta
Cruciplaccolithus neohelis	CCMP298	AJ246262	Chromalveolata	Haptophyta

Oolithus fragilis	ALGO AS641	AM491026	Chromalveolata	Haptophyta
Calcidiscus leptoporus	AS31	AJ544116	Chromalveolata	Haptophyta
Umbilicosphaera foliosa	ESP6M1	AJ544119	Chromalveolata	Haptophyta
Isochrysis litoralis	ALGO HAP18	AM490996	Chromalveolata	Haptophyta
Pseudoisochrysis paradoxa	CAP949/1	AM490999	Chromalveolata	Haptophyta
Dicrateria sp.	ALGO HAP49	AM490997	Chromalveolata	Haptophyta
Chrysotila lamellosa	ALGO HAP17	AM490998	Chromalveolata	Haptophyta
Isochrysis sp. zhangjiangensis		DQ075203	Chromalveolata	Haptophyta
Isochrysis sp.	CCAP 927/14	DQ079859	Chromalveolata	Haptophyta
Pleurochrysis carterae	HAP1	AJ544120	Chromalveolata	Haptophyta
Pleurochrysis dentata	HAP6	AJ544121	Chromalveolata	Haptophyta
Pleurochrysis sp	MBIC10443	AB183596	Chromalveolata	Haptophyta
Pleurochrysis gayraliae	ALGO HAP10	AM490972	Chromalveolata	Haptophyta
Pleurochrysis roscoffensis	ALGO HAP32	AM490974	Chromalveolata	Haptophyta
Hymenomonas globosa	ALGO HAP30	AM490981	Chromalveolata	Haptophyta
Hymenomonas coronata	ALGO HAP58 bis	AM490982	Chromalveolata	Haptophyta
Isochrysis sp.	MBIC10557	AB183617	Chromalveolata	Haptophyta
Jomonlithus littoralis	ALGO Je5	AM490979	Chromalveolata	Haptophyta
Ochrosphaera verrucosa	ALGO HAP82	AM490980	Chromalveolata	Haptophyta
Ochrosphaera sp.	MBIC10476	AB183604	Chromalveolata	Haptophyta
Isochrysis sp.	MBIC10464	AB058347	Chromalveolata	Haptophyta
Chrysochromulina scutellum	G7	AJ246274	Chromalveolata	Haptophyta
Prymnesiophyte symbiont 1		AF166377	Chromalveolata	Haptophyta
Chrysochromulina acantha	T20	AJ246278	Chromalveolata	Haptophyta
Chrysochromulina	J10	AJ246273	Chromalveolata	Haptophyta
Chrysochromulina sp.	MBIC10513	AB199882	Chromalveolata	Haptophyta
Chrysochromulina sp.	UIO TH2	AM491020	Chromalveolata	Haptophyta
Chrysochromulina rotalis	UIO P16	AM491025	Chromalveolata	Haptophyta
Chrysochromulina parva	CCMP 291	AM491019	Chromalveolata	Haptophyta
Chrysochromulina leadbeateri	UIO ERIK	AM491017	Chromalveolata	Haptophyta
Chrysochromulina sp.	NIES-1333	DQ980478	Chromalveolata	Haptophyta
Chrysochromulina simplex	UIO D4	AM491021	Chromalveolata	Haptophyta
Braarudosphaera bigelowii	TP05-6-b	AB250785	Chromalveolata	Haptophyta
Chrysochromulina parkeae	Kawachi, Japan	AM490994	Chromalveolata	Haptophyta
Braarudosphaera bigelowii	TP05-6-a	AB250784	Chromalveolata	Haptophyta
Imantonia rotunda	ALGO HAP23	AM491014	Chromalveolata	Haptophyta
Imantonia sp.	CCMP 1404	AM491015	Chromalveolata	Haptophyta
Imantonia sp.	MBIC10500	AB183606	Chromalveolata	Haptophyta
Chrysochromulina hirta	1Y	AJ246272	Chromalveolata	Haptophyta
Chrysochromulina cf.	CCMP 284	AM491011	Chromalveolata	Haptophyta
Chrysochromulina fragaria	UIO S19	AM491013	Chromalveolata	Haptophyta
Chrysochromulina brevifilum	PML 143	AM491012	Chromalveolata	Haptophyta
Prymnesium sp.	UIO 133	AM779755	Chromalveolata	Haptophyta
Chrysochromulina kappa	EN3	AJ246271	Chromalveolata	Haptophyta
Chrysochromulina minor	PLY 304	AM491010	Chromalveolata	Haptophyta
Chrysochromulina cf. polylepis	PCC200	AJ004868	Chromalveolata	Haptophyta
Prymnesium patelliferum		L34671	Chromalveolata	Haptophyta
Prymnesium nemamethecum		AJ246268	Chromalveolata	Haptophyta
Prymnesium sp.	MBIC10533	AB183612	Chromalveolata	Haptophyta
Prymnesium parvum	K081	AJ246269	Chromalveolata	Haptophyta
Prorocentrum minimum		EF017804	Chromalveolata	Haptophyta
Prymnesium zebrinum	ALGO HAP29	AM491001	Chromalveolata	Haptophyta
Prymensium sp.	ALGO HAP Pm	AM491002	Chromalveolata	Haptophyta
Platychrysis simplex	ALGO HAP51bis	AM491028	Chromalveolata	Haptophyta
Platychrysis pienaarii	ALGO HAP50 bis	AM491027	Chromalveolata	Haptophyta
Prymnesium faveolatum	ALGO HAP79	AM491005	Chromalveolata	Haptophyta
Prymensium sp.	ALGO HAPPM	AM491006	Chromalveolata	Haptophyta
Prymnesium annuliferum	ALGO HAP47	AM491007	Chromalveolata	Haptophyta
Prymnesium calathiferum	CCMP 707	AM491008	Chromalveolata	Haptophyta

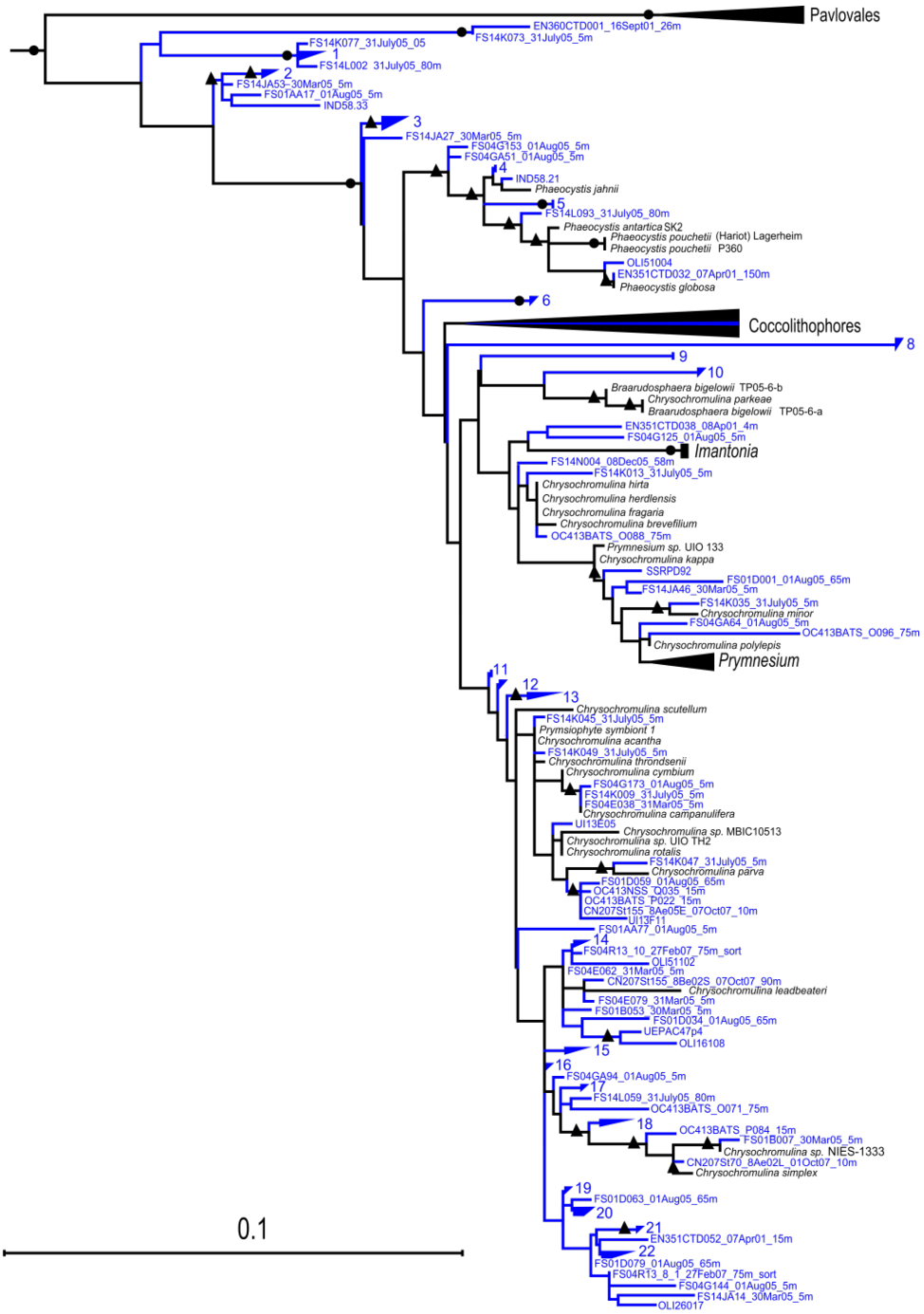
IND58.33	EU561900
OLI11007	AJ402346
OLI51080	AF107091
OLI26047	AF107085
IND58.21	EU561892
OLI51004	AF107086
OLI26041	AF107084
OLI51050	AF107088
SA2_1B5	EF527116
F01N5	EF173004
OLI51059	AF107089
OLI51033	AF107087
Q2G11N10	EF172980
SSRPD92	EF172993
IND70.03	EU562003
UI13E05	EU446347
UI13F11	EU446348
MB07.32	EF539132
UEPAC47p4	DQ369019
OLI16108	AF107082
OLI16029	AF107080
B04N10	EF172966
BL010625.10	AY426921
IND31.140	EU561805
N10E02	EF172967
OLI26017	AF107083

Table S3.3. Name of sequences present in collapsed clades in Figure 3.2. Sequences retrieved from size-fractioned (<2 or <3 μm) samples or flow sorts were collapsed after phylogenetic analysis and when all sequences within the clade had ≥ 99% identity. Names in *Italics* denote previously published sequences while other sequences are from this study.

Group number	Sequence name
1	FS01AA64_01Aug05_5m FS01AA86_01Aug05_5m FS14K020_31July05_5m
2	FS01D004_01Aug05_65m FS14JA13_30Mar05_5m
3	FS14L039_31July05_80m FS14I027_30Mar05_70m FS14M001_08Dec05_58m OC413BATS_O069_75m OC413BATS_P038_15m CN207St155_8Be04F_07Oct07_90m CN207St70_BBe08M_01Oct07_10m *OLI11007* *OLI16010* *OLI26047* *OLI51080*
4	OC413BATS_P009_15m EN351CTD040_40_09Apr01_4m EN351CTD040_09Apr01_20m
5	FS01B073_30Mar05_5m FS04E091_31Mar05_5m FS14JA52_30Mar05_5m OC413BATS_P053_15m
6	*OLI26041* *OLI51050*
7	FS04E051_31Mar05_5m FS14K029_31July05_5m
8	FS04R13_7_27Feb07_75m_sort FS04R14_1_27Feb07_75m_sort FS08L1_2_27Feb07_141m_sort FS14L086_31July05_80m FS14M077_08Dec05_58m *F01N5*
9	*OLI51059* *OLI51033*
10	OC413BATS_P036_15m *Q2G11N10*
11	FS04GA78_01Aug05_5m OC413BATS_P070_15m
12	FS14K084_31July05_5m *IND70.03*

13	FS01D092_01Aug05_65m OC413NSS_Q058_15m
14	FS01D027_01Aug05_65m FS04GA50_01Aug05_5m OC413BATS_P003_15m CN207St155_8Ae02Y_07Oct07_90m *MB07.32*
15	FS01B058_30Mar05_5m FS04H103_01Aug05_89m FS14L014_31July05_80m FS14M055_08Dec05_58m OC413NSS_Q042_15m OC413NSS_R062_70m *OLI16029*
16	FS01D093_01Aug05_65m FS04E093_31Mar05_5m FS04G183_01Aug05_5m OC413NSS_Q003_15m OC413NSS_Q028_15m
17	FS01D024_01Aug05_65m *B04N10*
18	FS01D021_01Aug05_65m FS01C021_30Mar05_70m CN207St155_8Ae02U_07Oct07_10m EN351CTD040_16_09Apr01_4m
19	OC413BATS_P088_15m *BL010625.10* *IND31.140*
20	FS01D058_01Aug05_65m *N10E02*
21	FS14L079_31July05_80m OC413BATS_O028_75m
22	FS01AA62_01Aug05_5m FS04GA79_01Aug05_5m

Figure S3.1. Maximum likelihood (ML) reconstruction of pico-prymnesiophyte SSU rRNA gene sequences from Fig. 2. Here all environmental clone names are shown, except for 99% identity Groups 1-22, for which clone IDs are given in Supplementary Information Section I Table S3.3. Note that all sequences within each 99% identity Group were used in the alignment (single representative sequences from each clone library or previously published data) and that the clades were collapsed after the tree was constructed for visualization purposes. Identity levels were determined using the original sequence (not from a masked alignment).



Supplementary Information Section II

Dynamics and Growth Rates in the Florida Straits and Sargasso Sea

The Florida Straits undergo a weak winter mixing and subsequently a long period of stratification through most of the rest of the year, much akin to that seen in the Sargasso Sea at BATS (Steinberg et al., 2001). In the Florida Straits, pico-prymnesiophytes ranged up to 1.2 x 10^3 cells ml^{-1} (December, Station 01) at the surface and 6.0 x 10^2 cells ml^{-1} (September, Station14) in the DCM but at times were below 100 cells ml^{-1} (e.g. May, Station 04 Surface, July/August Station 14 DCM; Supplementary Information Section II Figure S3.2). The majority of prymnesiophytes were picoplanktonic in size (Supplementary Information Section II Figure S3.2 and S3.3).

In the Sargasso Sea, cells were enumerated in samples collected from CTD profiles performed over four day (BATS) and three day (NSS) intervals, with dilution experiments being conducted on intervening days. Due to technical reasons, data from surface samples for CTD004 (BATS) and CTD056 (NSS) are not reported and we used samples from 10 m depth instead. Pico-prymnesiophyte cells binned as < 3 μm contributed substantially to the total number of pico-prymnesiophytes (Supplementary Information Section II Figure S3.4). These cells represented from 36% up to 68% of prymnesiophytes at BATS and between 18% and 62% at the NSS station. Other pico-prymnesiophytes fell within the 3-10 μm size fraction.

Specific cell sizing showed that prymnesiophytes formed two major size classes in the Sargasso Sea: those cells binned as <3 μm averaged 1.9 ± 0.4 μm x 2.1 ± 0.3 μm (n=89) and those in the 3-10 μm bin averaged 2.8 ± 0.6 μm x 3.4 ± 0.5 μm (n=161). Together these two groups composed 98 ± 2% of all prymnesiophytes detected. Prymnesiophytes

in the >10 μm were contributed 0-3 % of cells and were not considered further. Overall, pico-prymnesiophyte cell concentrations ranging from 177 ± 116 cells ml^{-1} (CTD056, DCM) to 872 ± 45 cells ml^{-1} (CTD056, surface) in the Sargasso Sea. Differences between mean cell concentrations were not significant between the two sites ($p = 1.0$). At BATS, average surface abundance of pico-prymnesiophytes were almost identical (536 ± 231 cells ml^{-1}) to those at DCM (539 ± 224 cells ml^{-1}). Nevertheless, these trends were not observed for individual CTD casts; numbers of cells were significantly higher at the surface than at the DCM for CTD004 ($p < 0.05$), the opposite was seen for CTD029 ($p < 0.01$). At the NSS station, pico-prymnesiophytes were more abundant at the surface (768 ± 129 cells ml^{-1}) than at the DCM.

Prochlorococcus and *Synechococcus* populations were also enumerated for both stations. In the upper water column, *Prochlorococcus* tended to be smaller and dimmer (i.e. containing less chlorophyll per cell) than at depth. Concentrations determined by FCM for *Prochlorococcus* in the surface waters at BATS therefore likely reflect underestimates, given that they were partially below the detection limit for chlorophyll on the InFlux flow cytometer. For our global data point (Figure 3.3) from this site we used a *Prochlorococcus* concentration from lower in the water column in order to compensate for this issue. *Prochlorococcus* clearly dominated in terms of abundance at both the surface and DCM in the Florida Straits, from March through December (Supplementary Information Section II Figure S3.5), ranging from 2.5 $x10^4$ cells ml^{-1} (June, Station 14) to 1.4 x 10^5 cells ml^{-1} (December, Station 14) at the surface and from 7.8 x 10^3 cells ml^{-1} (December, Station 01) to 1.4 x 10^5 cells ml^{-1} (June, Station 01). In the Sargasso Sea, at the NSS station, *Prochlorococcus* dominated throughout the water column, with a

maximum of 7.5 x 10^4 cells ml^{-1} at 15 m depth (Supplementary Information Section II Figure S3.6). Concentrations of *Prochlorococcus* were higher at the surface (above 40 m) and decreased below 40 m. This trend with depth was not observed for BATS data, where the highest cell concentrations were detected at 99 m (CTD004, DCM, 7.7 x 10^4 cells ml^{-1}) and 65 m (CTD029, above the DCM that was at 85 m, 1.2 x 10^5 cells ml^{-1}). *Prochlorococcus* abundance declined rapidly below 125 m for both stations (< 1.2 x 10^3 cells ml^{-1}). *Synechococcus* concentrations in the NSS were relatively constant above 80 m (31.40 ± 0.03 x 10^4 cells ml^{-1} for CTD056 and 21.5 ± 0.4 x 10^4 cells ml^{-1} for CTD081) and were consistently lower than those of *Prochlorococcus*. At BATS, average abundance above 20 m (1.5 ± 0.2 x 10^4 cells ml^{-1}) were comparable to those of *Prochlorococcus* but lower deeper in the water column.

Small eukaryotes were detected in all FCM samples in concentrations ranging from ~4 cells ml^{-1}, at the base of the euphotic zone, to 8.4 x 10^3 cells ml^{-1}. Overall, the vertical distribution of small eukaryotes was similar at BATS and the NSS. Concentrations were relatively homogeneously throughout the first 65 m of the water column; maximum concentrations occurred around the DCM (between 80 and 99 m). A sharp decrease in abundance was observed below 125 m. Comparison between the sites revealed that surface abundance of small eukaryotes was on average greater at the NSS station (2.3 x 10^3 ± 3.4 x 10^2 cells ml^{-1}) than at BATS (1.2 x 10^3 ± 1.3 x 10^2 cells ml^{-1}). However, at the DCM, the average cell concentrations were comparable for BATS and the NSS station (around 5.4 x 10^3 ± 4.2 x 10^3 cells ml^{-1} at BATS, 5.3 x 10^3 ± 1.2 x 10^3 at the NSS station). Interestingly, both the lowest and highest concentrations of small eukaryotes at the DCM were detected at BATS over the 4 day interval, indicating a dynamic system, likely

reflecting the progression of stratification that occurs at this time of year (Steinberg et al., 2001). Overall, pico-prymnesiophytes (detected by FISH) represented from 27% (CTD029) to 69% (CTD004) of the abundance of small eukaryotes recorded by FCM at the surface and between 4% (CTD056) and 14% (CDT004) at the DCM.

Picophytoplankton biomass at the NSS station was dominated by *Prochlorococcus* at most depths above 80 m (between 1.8 μg C l^{-1} and 3.0 μg C l^{-1}; Supplementary Information Section II Figure S3.6). At this site, average pico-prymnesiophytes biomass was higher at the surface than at the DCM (1.7 μg C l^{-1} ± 0.6 μg C l^{-1} and 0.8 ± 0.4 μg C l^{-1}, respectively). 'Non-prymnesiophyte' picoeukaryotes and *Synechococcus* also contributed significantly to the picophytoplankton biomass at the same depths (between 0.7 μg C l^{-1} and 1.5 μg C l^{-1} for the former and 1.4 μg C l^{-1} and 2.6 μg C l^{-1} for the latter). 'Non-prymnesiophyte' eukaryote biomass peaked at the DCM, reaching a maximum of 2.6 μg C l^{-1} for CTD081. At BATS, surface picophytoplankton biomass was lower than in the NSS. Generally (date dependent), maximum biomass was reached at the DCM (85 m, 3.0 μg C l^{-1}) or just above (65 m, 4.8 μg C l^{-1}) at BATS for most groups.

Dilution experiments were performed at these stations to determine the growth and grazing mortality rates of pico-prymnesiophytes (by FISH) and the other picophytoplankton groups (by FCM). Two experiments were performed at BATS, Exp. 1 (75 m) and Exp. 2 (15 m) and two at the NSS station, Exp. 3 (15 m) and Exp. 4 (70 m). Pico-prymnesiophyte abundances were 324 ± 136 cells ml^{-1} and 238 ± 94 cells ml^{-1} in Exp. 1 and Exp. 2, respectively. These concentrations were slightly lower than values at the surface and DCM of the day previous to each experiment (see Supplementary Information Section II Figure S3.4), in part because the DCM was deeper in the water

column, i.e. Exp. 1 was conducted in a region of the water column with lower abundance than the DCM. For Exp. 3 and Exp. 4 cell concentrations were 448 ± 144 cells ml^{-1} and 651 ± 282 cells ml^{-1}, respectively.

An ANOVA was performed to test the significance of the regression used to analyze dilution experiments data, only rates from statistically significant regressions ($p < 0.1$) were used. Low abundance, which was magnified in dilution treatments, made it difficult to enumerate sufficient numbers of cells in diluted bottles, particularly at BATS. The ANOVA results led to the BATS experimental data for pico-prymnesiophytes and small eukaryotes being discarded. At the NSS station, pico-prymnesiophyte growth and grazing mortality rates were higher at the surface (1.12 day^{-1}, 1.41 day^{-1}, for growth and grazing mortality rates respectively, r^2=0.87, p = 0.06) than deeper in the water column (70 m; 0.29 day^{-1}, 0.70 day^{-1} for growth and grazing mortality rates respectively r^2=0.73, p = 0.06). For small eukaryotes, the opposite trend was observed, with higher growth and grazing mortality rates at depth (0.51 day^{-1},0.74 day^{-1}, for growth and grazing mortality rates respectively, r^2=078, $p < 0.0001$) than at the surface (0.22 day^{-1}, 0.29 day^{-1}, for growth and grazing mortality rates respectively r^2=0.41, p = 0.06). Quantitative PCR data indicates that the DCM at this station was dominated by the prasinophyte *Ostreococcus* (Cuvelier and Worden, unpubl.). Clearly a greater number of experiments are needed to ascertain the frequency with which the growth rates recorded here are encountered. Using rate and biomass data for pico-prymnesiophytes, we estimated that the amount of primary production at the NSS site was 1.1 µg C l^{-1} day^{-1} at the surface, or 2.4 µg C l^{-1} d^{-1} if production is considered without the effect of grazing, almost 4 fold more than for other picoeukaryotes (0.27 µg C l^{-1} day^{-1}, or 0.30 µg C l^{-1} d^{-1} if production is considered

without the effect of grazing). Again, these roles were reversed at 70 m where pico-prymnesiophyte primary productions was 0.3 μg C l^{-1} day^{-1} (or 0.5 μg C l^{-1} d^{-1} if production is considered production was observed without the effect of grazing) and 'non-prymnesiophyte' picoeukaryotes were estimated to produce 1.8 μg C l^{-1} day^{-1} (2.6 μg C l^{-1} d^{-1} if production is considered based on absence of grazing).

FISH and HPLC data was compared for the Florida Straits time series. Prymnesiophytes formed a large fraction of the picophytoplankton biomass based on HPLC (Supplementary Information Section II Figure S3.7). FISH counts agreed relatively well with HPLC at the surface, but indicate that HPLC overestimated biomass at the DCM significantly ($p < 0.0001$; Supplementary Information Section II Figures S3.7-S3.9). Dinoflagellates have been seen frequently in the tropical open ocean, although this data is unpublished (Chavez et al., unpubl.). Should these dinoflagellates contain Prymnesiophyceae-type pigmentation, as has been seen by others (Mackey et al., 1996; Latasa et al., 2001), they could contribute to HPLC overestimation of prymnesiophytes. In addition to possible issues with chlorophyll to specific pigment ratio algorithms other issues might result in the discrepancy between DCM data from these two methods. For instance, the FISH-based identification might have missed some Prymnesiophyceae at the DCM, although there was no statistical difference between these FISH and microscopy-based counts when North Atlantic data was compared for samples from multiple locations (averages of: 593 ± 108 (SE) cell ml^{-1}, characteristic-based microscopy vs. 500 ± 61 (SE) cell ml^{-1}, FISH). Sizing of PRYM02 hybridized cells

demonstrated that almost all prymnesiophytes were <3 μm (90% of cells at the surface and 87% of cells at the DCM, for all Florida Straits data; Supplementary Information Section II Figure S3.3).

Figure S3.2. Seasonal pico-prymnesiophyte cell concentrations in the Florida Straits. Data are shown for a 2005 time-series sampling at three Florida Straits stations over the course of the year. Concentrations were determined by TSA-FISH using a prymnesiophyte specific probe. Dark turquoise, <3 μm cells; light turquoise, 3-10 μm cells, measured cells this range averaged 2.8 μm x 3.4 μm (see Table S3.3); black, total numbers of cells including the <3 μm size fraction, 3-10 μm size fraction and >10 μm size fraction. The green line represents the *in vivo* fluorescence (unprocessed data) signature from the rosette mounted fluorometer (not available for December 2005).

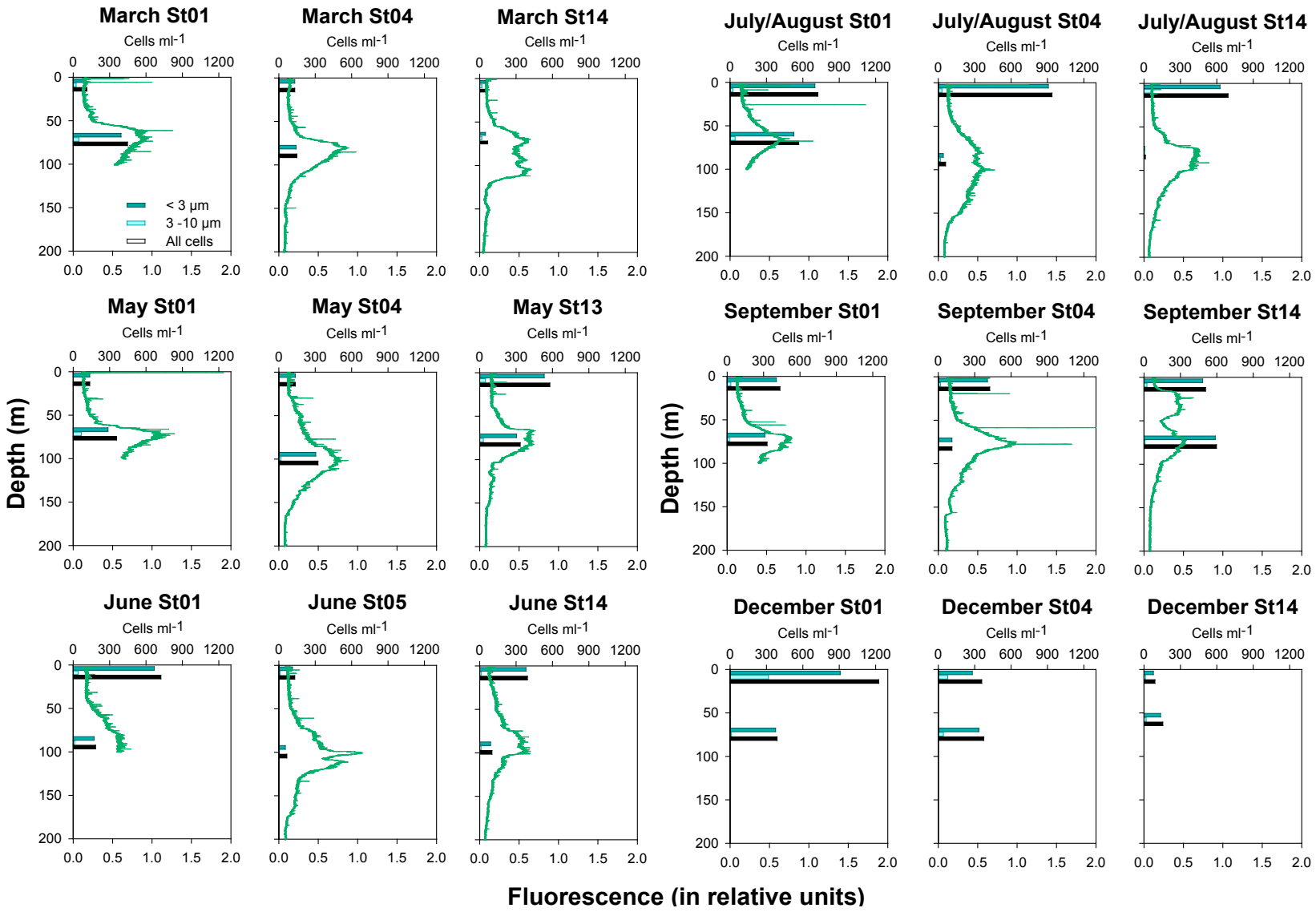

March St01
March St04
March St14
July/August St01
July/August St04
July/August St14
May St01
May St04
May St13
September St01
September St04
September St14
June St01
June St05
June St14
December St01
December St04
December St14
Cells ml-1
0 300 600 900 1200
0 50 100 150 200
0.0 0.5 1.0 1.5 2.0
< 3 µm
3 -10 µm
All cells
Depth (m)
Fluorescence (in relative units)

Figure S3.3. Percent of prymnesiophytes in each size fraction in the Florida Straits transect. Samples were taken on 6 cruises in 2005. (a) At the surface (5 m), 90% (9 SD) of the FISH probed prymnesiophytes were <3 µm (dark turquoise) and 9% (9 SD) were in the 3-10 µm range (light turquoise), while 0% (1.1 SD) were in the >10 µm size fraction (white). (b) At the DCM, 87% (12 SD) of the cells fell in the <3 µm size fraction and 12% (12 SD) within the 3-10 µm size range while 1% (2 SD) were in the >10 µm size fraction. Measured cells in the 3-10 µm range averaged 2.8 mm x 3.4 mm (see Table S3.3).

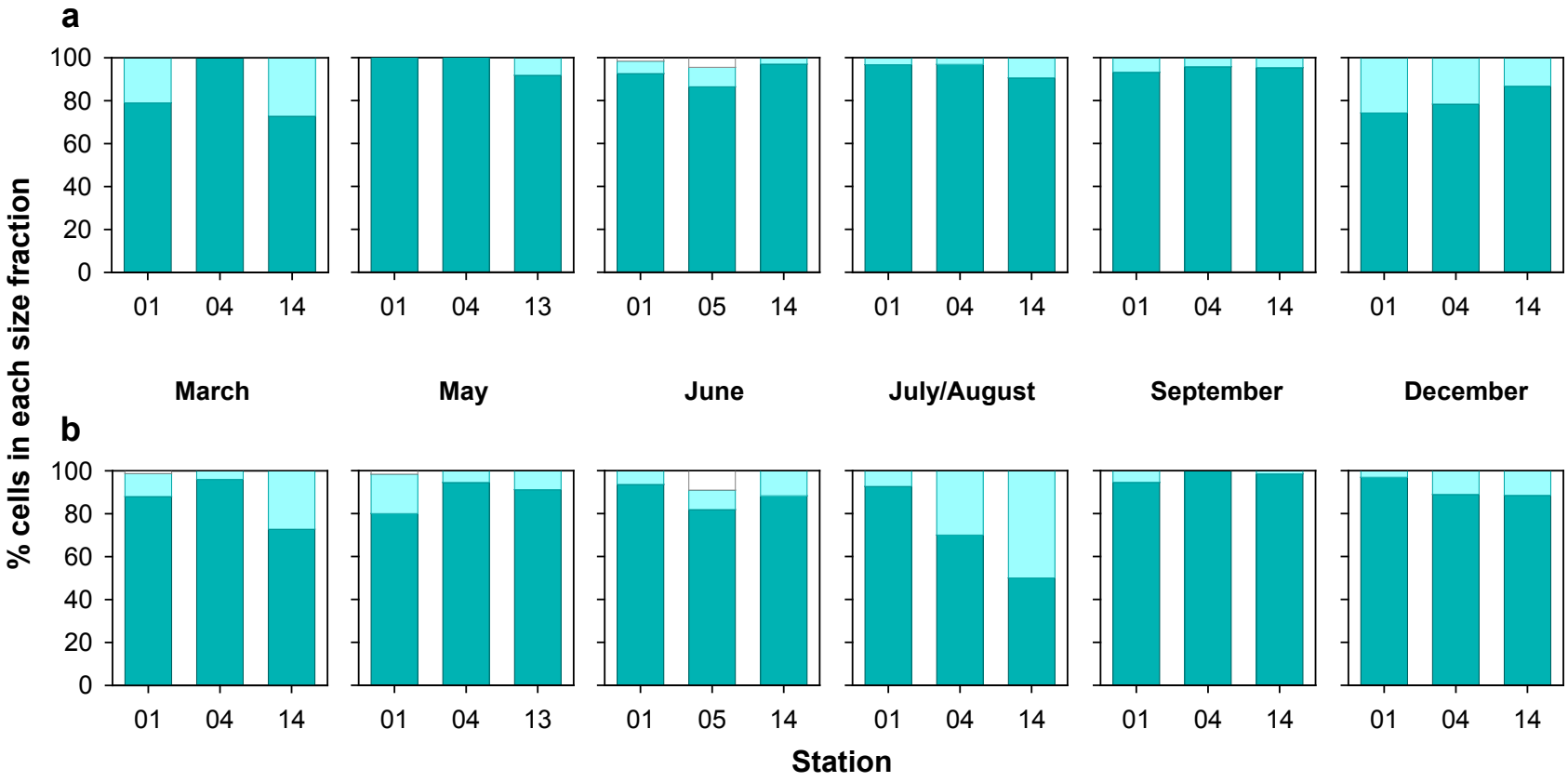


Figure S3.4. Pico-prymnesiophyte cell concentrations at the Sargasso Sea stations. Data are shown for two profiles (a) CTD004, 27 May 2005 and (b) and CTD029, 31 May 2005, both at BATS and for (c) CTD056, 4 June 2005 and (d) CTD081, 7 June 2005, both at the Northern Sargasso Sea station. Cell concentrations were determined by TSA-FISH using a prymnesiophyte specific FISH probe. Dark turquoise, <3 µm cells (average cell size: 1.9 ± 0.4 µm x 2.1 ± 0.3 µm); light turquoise, 3-10 µm cells (average cell size: 2.8 ± 0.6 µm x 3.4 ± 0.5 µm); black, total abundance (the sum of the different size categories). Error bars represent the standard deviation of two hybridizations. In the Northern Sargasso, the average abundance was statistically higher at the surface than at the DCM ($p < 0.02$); no statistical difference was detected at BATS ($p < 0.98$). The light green line represents the *in vivo* fluorescence signature from the rosette mounted fluorometer.

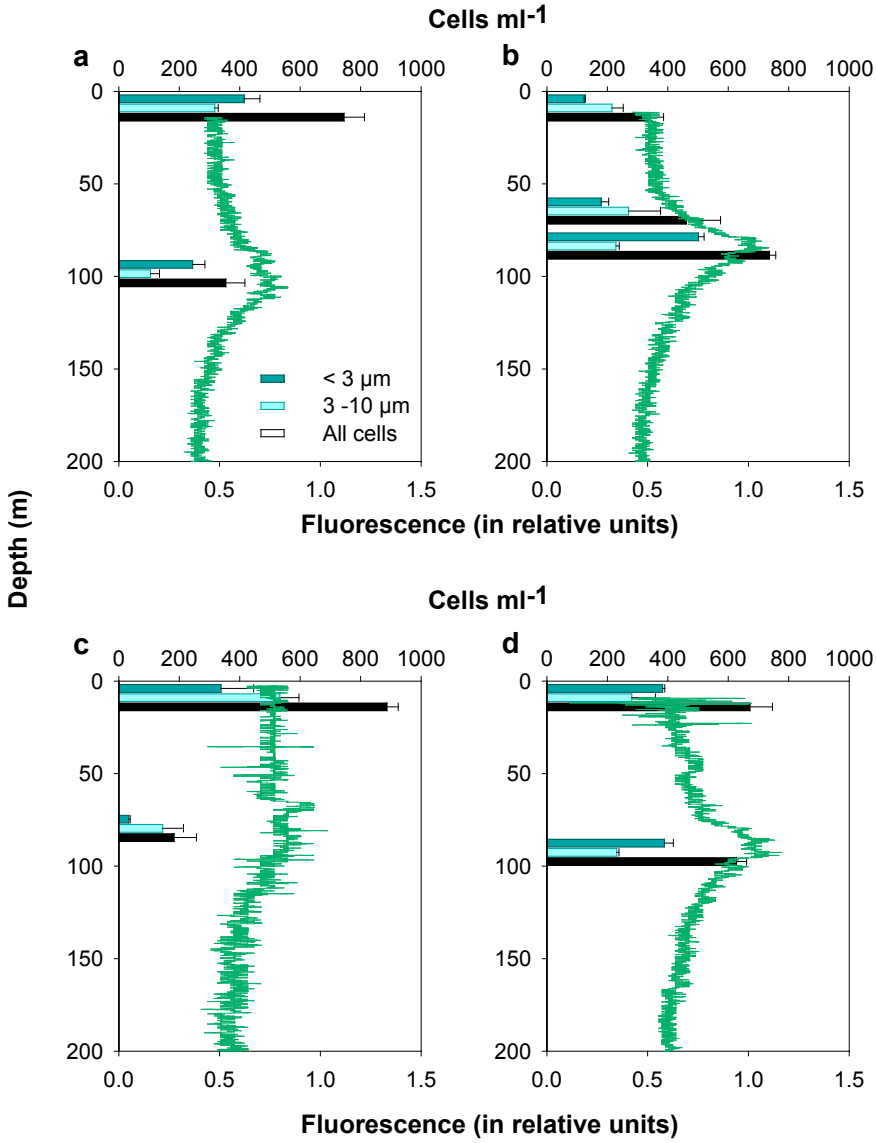


Figure S3.5. Picophytoplankton cell concentrations in the Florida Straits. Concentrations at (a) the surface (5 m) and (b) DCM (which varied in depth, falling between 69 and 100 m depending on date) of (purple) small eukaryotes, (black) *Synechococcus* and (grey) *Prochlorococcus* over the course of the year at three sampling sites (Station 01, 04/05 and 14). The core of the Gulf Stream, as represented by ADCP data shifted, hence Station 04 and 05 (fixed longitudes) represented the core depending on the season. Cell abundance determined by flow cytometry. Triangle indicates data points for which *Prochlorococcus* cells were partially offscale and therefore likely over-represent contributions of the other picophytoplankton groups.

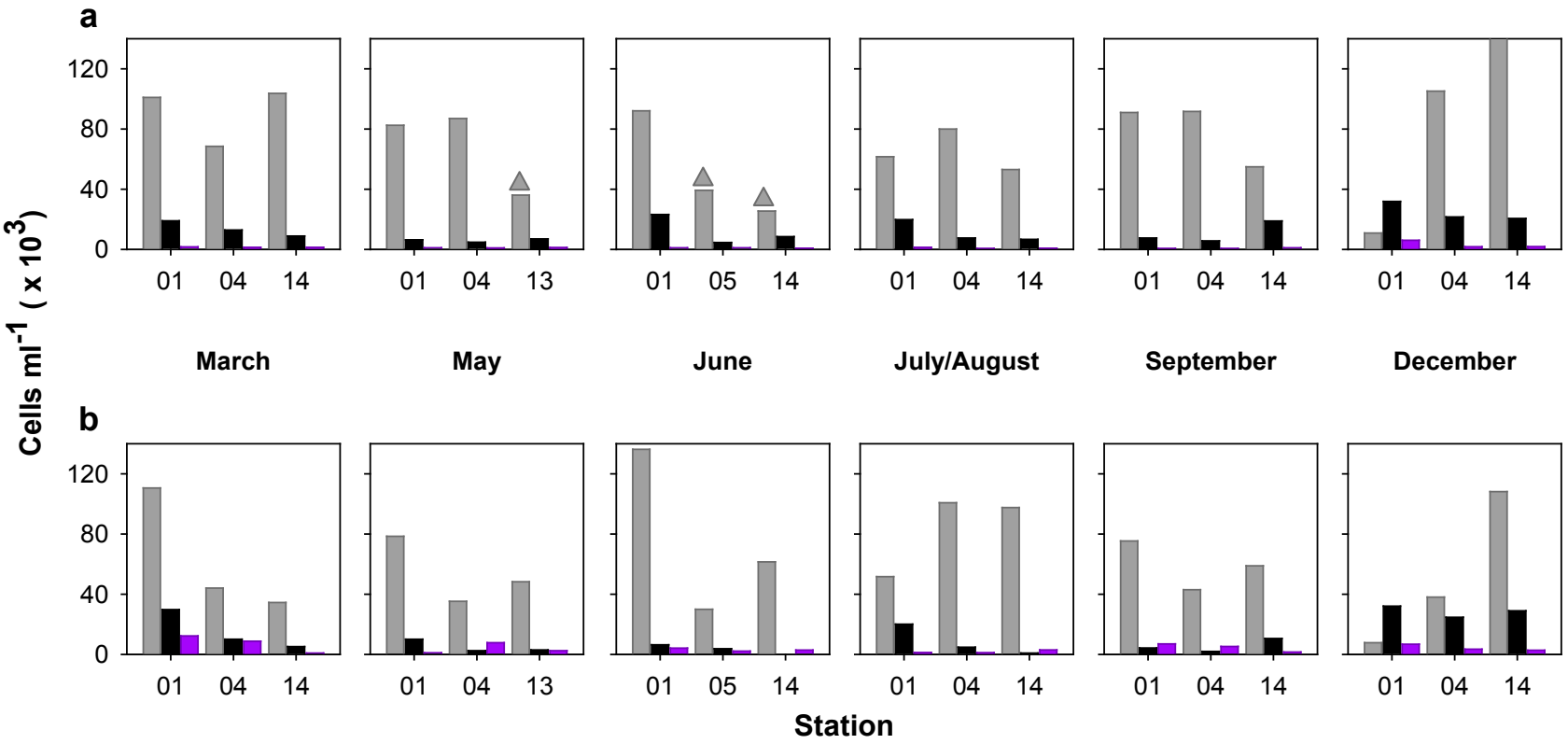


Figure S3.6. Cell concentrations and biomass of picophytoplankton at the Sargasso Sea stations. Samples were taken at BATS (CTD004, 27 May 2005 and CTD029, 31 May 2005) and at the Northern Sargasso Sea station (CTD056, 4 June 2005 and CTD081, 7 June 2005). Shown are (purple) small eukaryotes, (black) *Synechococcus* and (grey) *Prochlorococcus* enumerated by (a) flow cytometry. (b) Biomass for the same groups was estimated using the biovolume-based biomass conversion factors in Table S3.3.

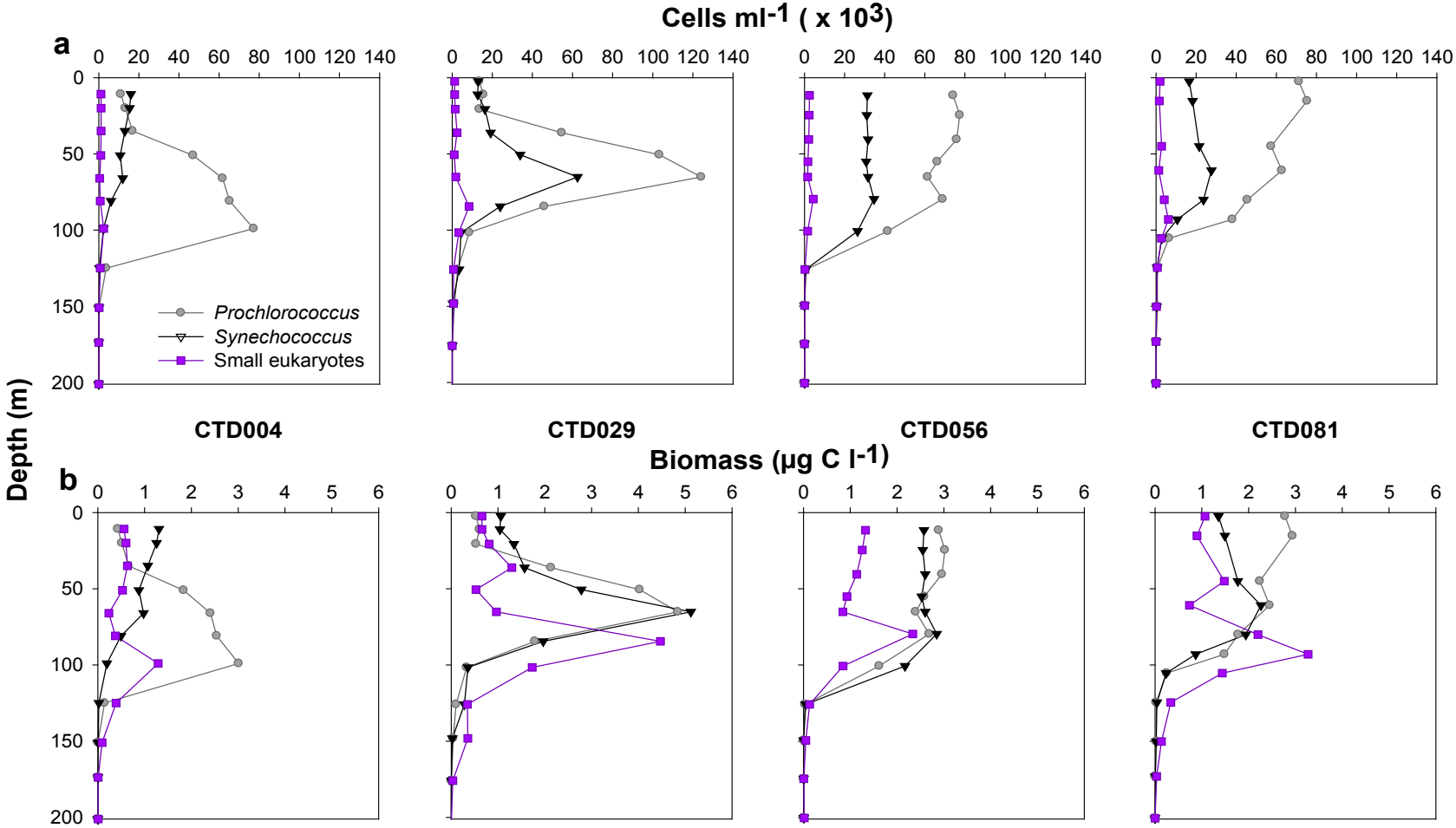


Figure S3.7. HPLC pico-prymnesiophyte contributions to total picophytoplankton biomass in the Florida Straits. Concentrations are shown for the (a) surface and (b) DCM. Percent biomass contributions over the 2005 year for (blue) prymnesiophytes, (purple) 'non-prymnesiophyte' eukaryotes as the sum of other eukaryotic taxa, (grey) *Prochlorococcus* and (black) *Synechococcus*.

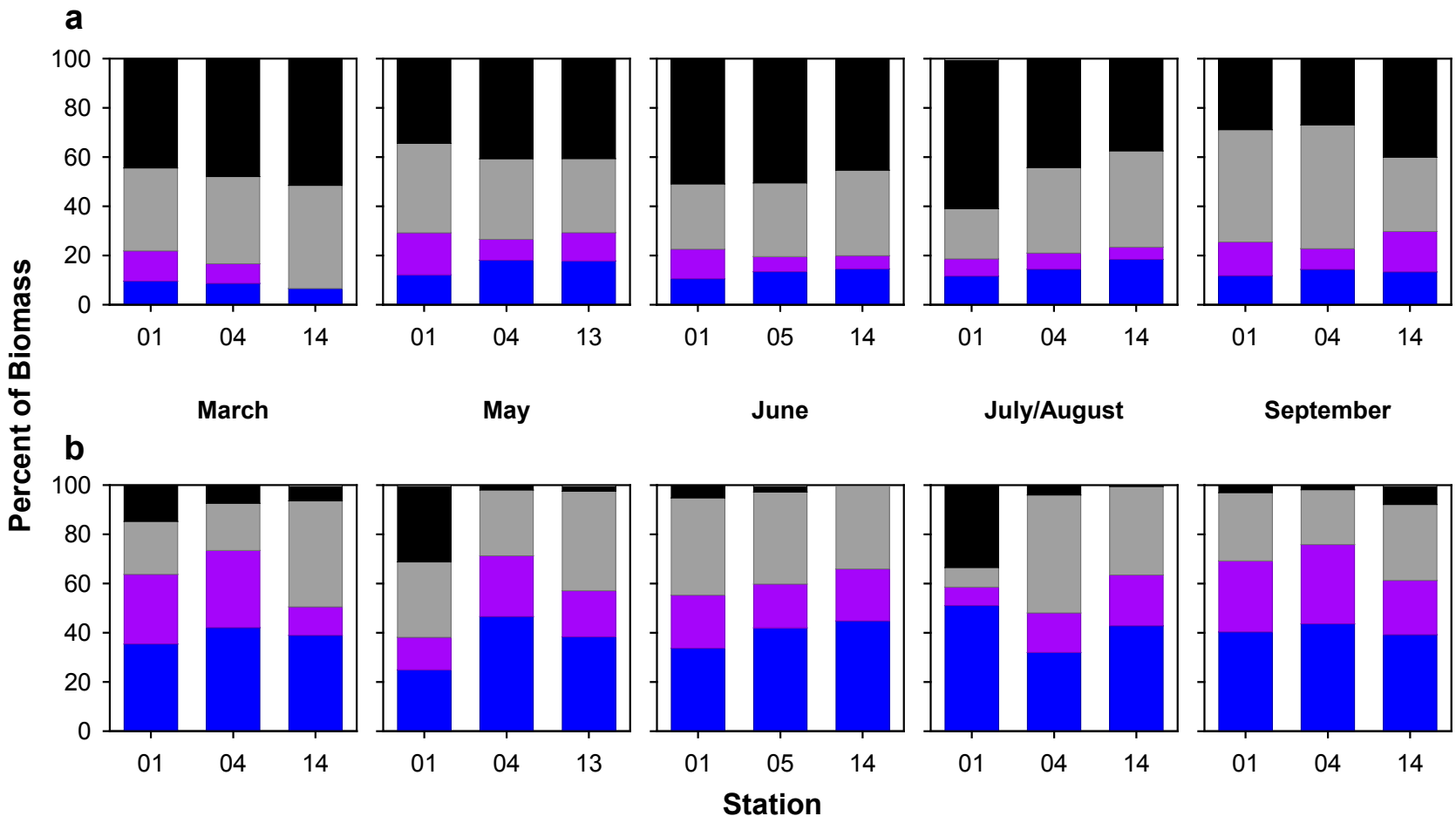


Figure S3.8. Pico-prymnesiophyte contributions to total picophytoplankton biomass using two methods. Biomass based on (hatched blue) direct counts and (blue) HPLC in the Florida Straits. Data are shown for (a) surface and (b) DCM samples over the course of the year. For direct count percentages, the abundance of pico-prymnesiophytes was determined by TSA-FISH and converted to carbon biomass using biovolumes, while overall picophytoplankton abundances were determined by FCM and converted to biomass using biovolume, abundance and a conversion factor (Worden et al. 2004). Note that the observed discrepancies at depth could be driven by issues with interpreting pigment measurements. Alternatively, it could indicate an underestimate of pico-prymnesiophyte contributions by FISH. Given the efficiency of the probe on control samples and the fact that comparison of two direct count approaches showed no statistical difference, this explanation seems unlikely. Over representation of other picophytoplankton (driving the percentage of pico-prymnesiophyte contributions down) by e.g., overestimation of carbon per cell for the other three groups, could also result in the observed discrepancy. Finally, the fact that large cells were not included in our FISH analysis, but were included in HPLC data could lead to differences. The majority of cells belonged to the pico-size fraction (Supplementary Information Section II Figure S3.3), however given the degree to which biovolume increases with increased radius results might still be influenced.

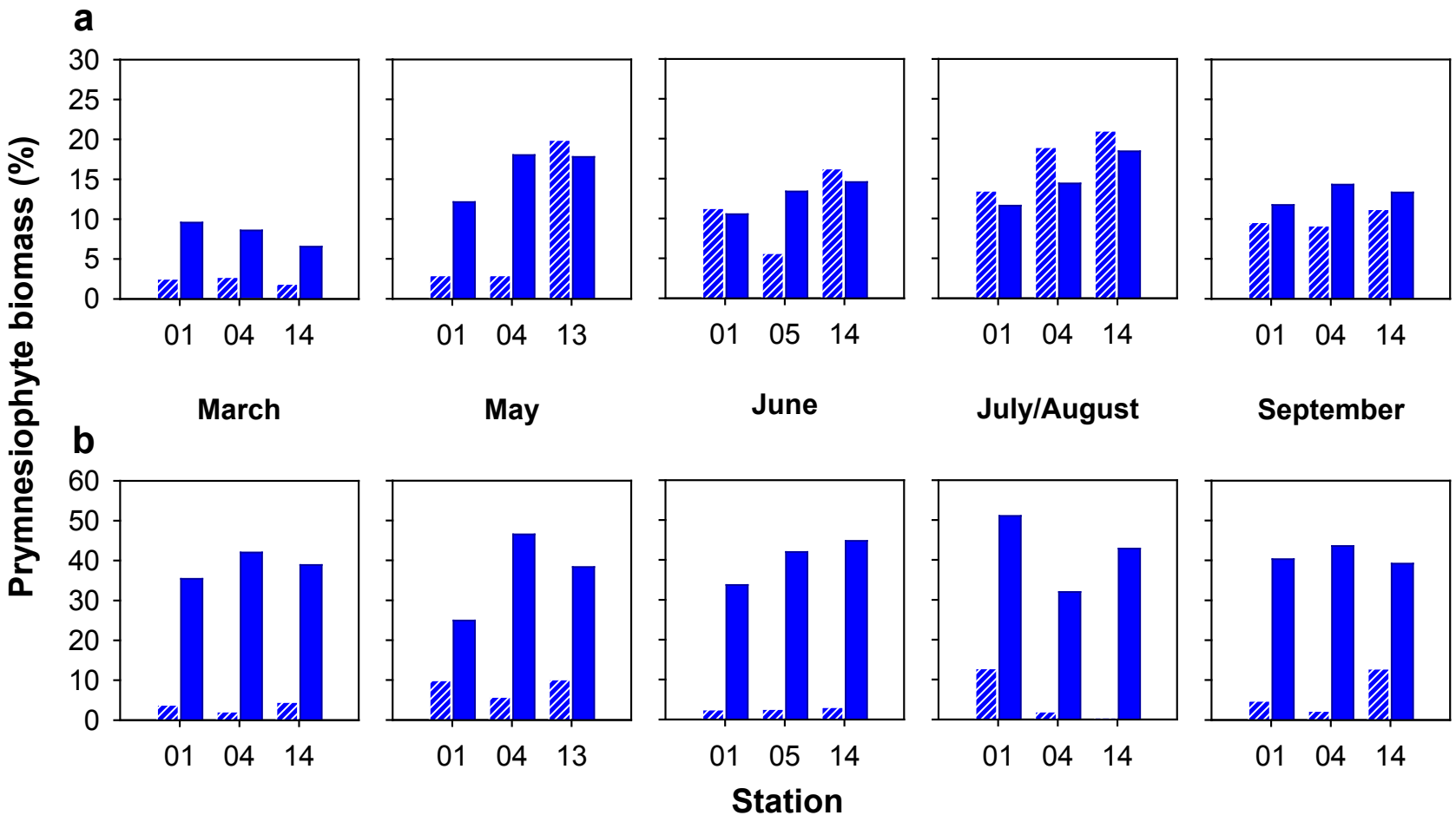
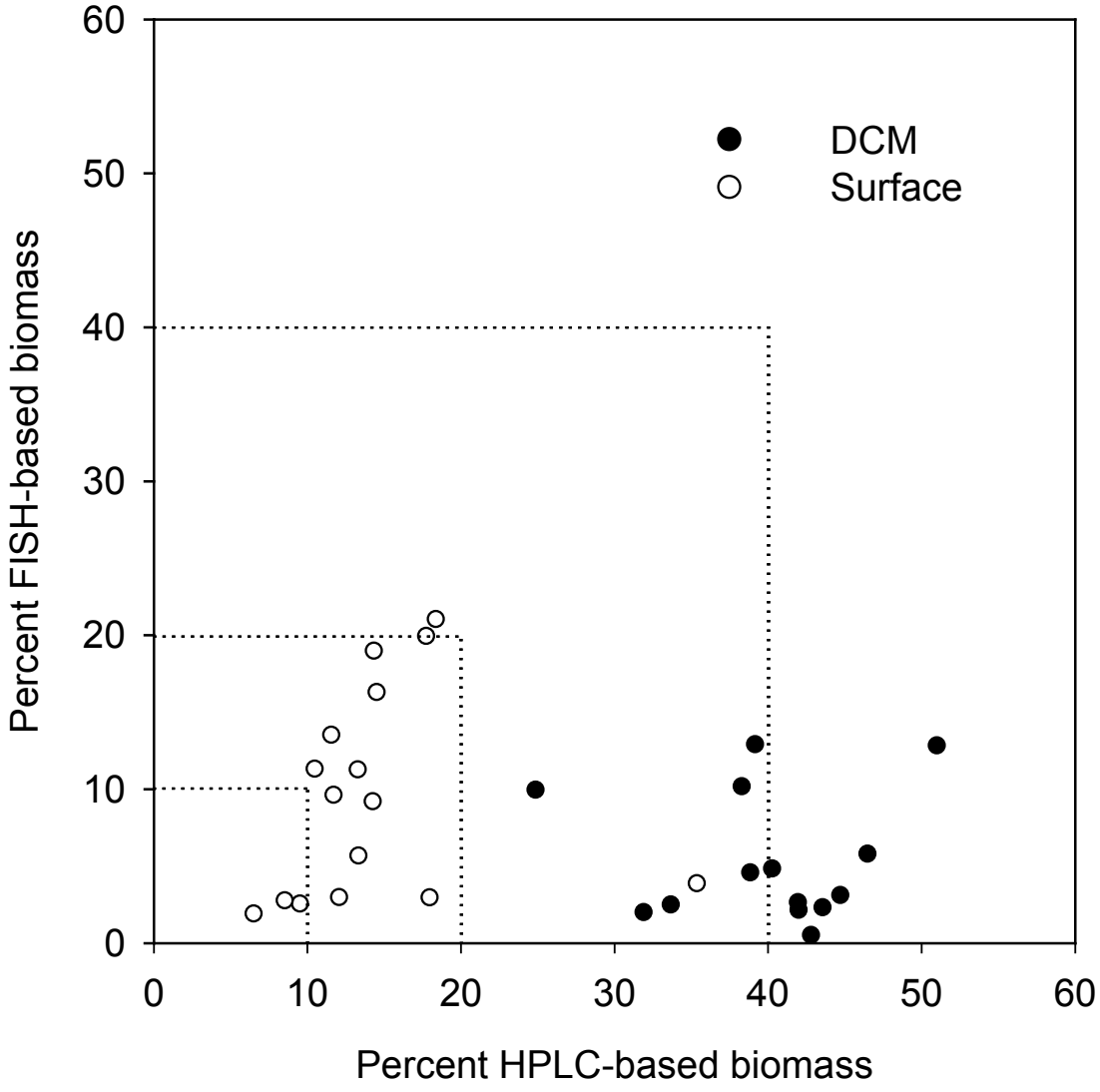

a
b
Prymnesiophyte biomass (%)
March
May
June
July/August
September
Station
01 04 14
01 04 13
01 05 14
01 04 14
01 04 14

Figure S3.9. Comparison of HPLC-based and FISH-based prymnesiophyte biomass contributions (%) to picophytoplankton biomass in the Florida Straits. At the DCM (black circles), prymnesiophyte biomass by HPLC appeared to overestimate their average contribution relative to the combined FISH (pico-prymnesiophytes) and FCM (other picophytoplankton groups) based-estimates from the 2005 time-series. Alternatively, the FISH and FCM based estimates may have underestimated prymnesiophyte contributions, or over estimated contributions by the other three picophytoplankton groups. At the surface (white circles), the relationship between these two data types was closer and no significant difference was detected. Only surface data was used to generate global biomass contribution data.



Acknowledgements

Andrew E. Allen, Adam Monier, Monique Messié, Susannah G. Tringe, Tanja Woyke, Rory M. Welsh, Thomas Ishoey, Brian J. Binder, Mikel Latasa, Cédric Guigand, Kurt R. Buck, Jason A. Hilton, Roger S. Lasken, Francisco P. Chavez and Alexandra Z. Worden contributed to this work and are listed as co-authors on a manuscript submitted that incorporates this chapter with metagenomic sequencing of the flow sorted samples. The submitted manuscript includes other co-authors, not included in the above list, who, along with some of the above, participated in analyses not presented here, these are Andrew E. Allen, John P. McCrow, Jae-Hyeok Lee, Chris L. DuPont, Mathangi Thiagarajan, Elisabet Caler, Betsy A. Read. This research was made possible by the captains and crews of the Research Vessels used, and facilitated by A. Engman and E. Demir as well as especially F. Not. R. Cowen provided several cruise opportunities. S.J. Giovannoni and R. Gausling provided manuscript comments, H. Wilcox, G. Weinstock, J. Heidelberg and R. Gausling contributed to 18S rDNA sequencing. Finally, J. Eisen and C. Langdon deserve thanks. Sequencing was performed under a CSP award to AZW and J. Eisen under the auspices of the BERP, US DOE Office of Science, and by the University of California, Lawrence Berkeley National Laboratory contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory contract No. DE-AC02-06NA25396. Some research was supported by DOE grant DE-FC02-02ER63453 and NSF-MCB-0732448 to AEA, much of the global phytoplankton data set by NOAA and the David and Lucile Packard Foundation grants to FPC and OC413 (dilution experiments) by

OCE-0241740 to BJB. This research was funded by NSF OCE-0836721, the David and Lucile Packard Foundation and a Gordon and Betty Moore Foundation Young Investigator Award to AZW.

Chapter 4:

Photoacclimation and adaptation in the picoeukaryotic prasinophyte *Micromonas*

Background

Ultraviolet (UV) radiation and high light (HL) can both negatively affect the phytoplankton growth, survival, pigmentation, metabolism, and photosynthesis (Xue et al., 2005). In turn this can impact primary production and have consequences for higher trophic levels that rely on that production. Penetration of UV radiation in the ocean depends on the wavelength, characteristics of the water and the geographic location (Castenholx and Garcia-Pichel, 2000). UV-B (280-320 nm) are absorbed more rapidly than UV-A (320-400 nm) but can still affect organisms in some regions. For example, in the Sargasso Sea, 10% of the incident 310 nm radiation (UV-B) present at the surface is still present at 20 m (Kirk, 1994). Other water bodies such as the Baltic Sea can absorb UV-B much more effectively, and radiation reaches only 70 cm (Kirk, 1994).

The consequences of HL and UV light exposure can be severe for organism health. If not repaired, damages induced by HL and UV radiation result in the decrease of photosynthetic capacities of algae, a process called photoinhibition. Under photoinhibition, the electron transport in the photosystem (II) decreases (Tevini et al., 1991). Furthermore, under excessive light, reactive oxygen species are produced in the chloroplast (Hutin et al., 2003). These can cause oxidative damage and irreversible effects to various components of the photosynthetic apparatus (Norén et al., 2003, Niyogi, 1999). In order to protect themselves against HL and UV radiation and reduce the negative effects of oxidative processes or prevent the formation of oxygen reactive

species, photosynthetic organisms have developed various photoprotection or acclimation mechanisms (Niyogi, 1999). These are tightly coupled with the light-harvesting complexes (LHC) that are essential for channeling light energy for use in photosynthesis.

Most photosynthetic organisms possess two photosystems (PS): PSI and PSII with LHCI and II as their respective LHC. In plants, the PSII antenna contains a permanent inner antenna (three monomeric minor proteins: CP29, CP26 and CP24) and a peripheral and potentially mobile major antenna (homologous and heterologous trimers of LHCII proteins: LHCB1, LHCB2 and LHCB3, Jansson, 1999). In prasinophyte algae, the structure of PSII appears to be very similar to plants with a few exceptions. There are several noticeable differences in pigment-protein complexes in *Mantoniella squamata*, the first representative of the class prasinophyte order Mamiellales to be explored. Studies have shown that only one unique LHC type, named LHCP ("P" for prasinophyte), was present in this organism. This is interesting because it is hypothesized that in ancestral algae a single LHC was associated with both PS; hence *Mantoniella* might represent such an ancestral state. However, the presence of LHCI genes and proteins was revealed by Six, Worden and colleagues (2005) in *Ostreococcus tauri* another member of the Mamiellales, in addition to the LHCP. In this study the authors returned to *M. squamata* and showed that the earlier work had simply "missed" LHCI proteins in that organism due to the use of overly specific antibodies. Therefore, even though the LHCP proteins were abundant LHC proteins in Mamiellales, LHCI proteins also formed a significant fraction of the LHC proteins. These findings raise questions about the specific function of LHCP and potential specific adaptations of prasinophytes to a range of light fields, intensities or variations. Even though it is likely that LHCP are

associated with PSII, the phylogenetic position of LHCP proteins outside the clade containing the LHCII polypeptides of plants and green alga, make the Mamiellales a unique model (Koziol et al., 2007). The complete genome sequence (Worden et al., 2009) of the picoeukaryotic green algae *Micromonas sp.* RCC299 (Mamiellales) possesses four *Lhcp* genes (*Lhcp1-4)*, one of which (*Lhcp2*) has seven copies (Worden et al., 2009). In this dissertation, the expression of *Lhcp* transcripts under HL and UV light is investigated.

Under light stress, algae have multiple responses strategies, such as the alteration the PS apparatus composition and the production of protective pigments and proteins (Norén et al., 2003, Salem and van Waasbergen, 2004). One example of proteins is the accumulation of light stress proteins from the early light-inducible proteins (ELIPs) family (Heddad and Adamska, 2002), a class of nuclear-encoded proteins structurally related to the LHC proteins and distant relatives of the chlorophyll *a/b* binding protein (Hutin et al., 2003, Montané and Kloppstech, 2000; Heddad et al., 2006). The first reports of ELIPs are associated with the early phase of greening in etiolated pea seedlings (Meyer and Kloppstech, 1984). Transcripts of these genes are also shown to transiently increase during the maturation stage of the chloroplast (Dolganov et al., 1995). In higher plants and green algae, ELIPs have three transmembrane domains, two of which have conserved LHC motifs (Adamska et al., 1992; Green and Kühlbrandt, 1995; Adamska, 1997), suggesting that they can bind to chlorophyll. It was therefore proposed that ELIPs are involved in the photoprotection by transiently binding to chlorophylls released in plants exposed to light stress (Montané and Kloppstech, 2000; Adamska et al., 2001).

Nevertheless, even though the regulation of the genes and proteins they encode is modulated by light and other stress signals, their exact functions and roles remain unclear (Harari-Steinberg et al., 2001).

Another type of stress-related proteins is the LHCSR (Peers et al., 2009, formerly known as LI818, Gagné and Guertin, 1992). The LHCSR polypeptide shows weak but significant homology to chlorophyll *a/b*-binding proteins of plants and algae (Gagné and Gertin, 1992, Savard et al., 1996). In addition, LHCSR are expected to have three transmembrane helices (Richard et al., 2000). Similarly to *Elip* genes, *Lhcsr* transcripts accumulate under conditions that cause photo-oxidative stress, such as excessive light, as well as CO_2 deprivation, sulfur and iron deprivation (Miura et al., 2004, Zhang et al., 2004, Naumann et al., 2007, Leford et al., 2004). Orthologues of this protein are present in all photosynthetic eukaryotes except vascular plants and red algae (Koziol et al., 2007, Peers et al., 2009). In the unicellular green algae *Chlamydomonas reinhardtii*, LHCSR is required for the thermal dissipation of excessive energy (or non-photochemical quenching), as mutants that lack LHCSR genes do not survive under shifts to HL (Peers et al., 2009). Cyanobacteria also express stress proteins. Under HL, cyanobacteria produce small polypeptides called high light-induced proteins (HLIPs), one helix proteins (OHP) with similarity to ELIPs (Dolganov et al., 1995, He et al., 2001, van Waasbergen et al., 2002, Salem and van Waasbergen, 2004). HLIPs may be involved in non-photochemical quenching and are necessary for the survival of cells. Mutants which do not possess any of these genes rapidly died under HL (Hsiao et al., 2004, van Waasbergen et al., 2002). *Micromonas* possess 2 *Ohp* genes and the expression of one will be explored in this work, in addition to the expression of a *Lhcsr* gene.

Material and Methods

Cultures for establishing growth vs. irradiance

Micromonas sp. strain RCC299 cells were grown in semi-continuous batch culture at 21 °C on a 14:10 light: dark (L:D) cycle in K medium (Keller et al., 1987) prepared with artificial seawater (see http://www.mbari.org/phyto-genome/Resources.html). Cultures were grown in 50 ml glass test tubes cleaned with 3.7% HCl. Triplicate cultures were grown at each approximate light level: 6, 55, 150, 240, 350, 470, 620, 750 μE m^{-2} s^{-1} (measured inside the test tube with a QSL-2101 light meter, Biospherical Instruments, San Diego, CA, USA). Low light levels were obtained by shading the tube with neutral density standard filters (Lee Filters, Burbank, CA, USA). Bulk fluorescence was measured daily at the same time of the L:D cycle using a 10-AU fluorometer (Turner Design, Sunnyvale, CA, USA) to determine growth rate. Cells were acclimated to each light level for 10 generations after which average growth rates was calculated.

Cultures for experimental manipulations

A derivative of *Micromonas sp.* RCC299, specifically RCC299 2.9 series (2.9 f21R1SDA1XA299.1.3.1.17b.49c), was used in all the following experiments. This strain's genome has been completely sequenced (Worden et al. 2009). Cultures were grown in K medium (Keller et al., 1987) made with artificial seawater (as outlined at above URL). The cells were grown in semi-continuous batch culture in mid-exponential phase from 20 June to 31 Aug 2007 (19.5 h experiments) and 20 June to 16 Aug 2008 (2.5 h experiments) at a light level of ~100 μE m^{-2} s^{-1} and at a temperature of 21.5 °C. Cells were grown on a 14:10 L:D cycle, with lights on at 6:00 and off at 20:00 during

summer 2007 (19.5h experiments), and on at 8:00 and off at 22:00 for summer 2008 (2.5h experiments). Cells were monitored daily by flow cytometry (FCM) using an Epics XL (Beckman Coulter, Brea, CA, USA see below) and growth rate was calculated. Cultures were transferred daily to reduce the concentration of cells to approximately 2 million cells ml^{-1}. The first experiment to be evaluated herein was performed after 9.1 generations and the average growth rates of the cells prior to this experiment was 1.14 day^{-1}. Every week or multiple times a week, a new experiment was performed. In order to ensure that cultures remained free of bacterial contamination, each week an aliquot was placed in a bacterial test media; rather than these returned negative results.

Experimental system setup

On day one of the onset of experimental manipulations, 1 l Erlenmeyer flasks were filled with 1 l of exponential phase culture within 30 minutes prior to the start of the experiment (t=0 h). All the flasks and culture bottles were handled in a laminar flow hood to avoid bacterial contamination. The top of all flasks and bottles were quickly flamed before and after pouring. Flasks were then placed into each water baths on a shaker and shaken constantly throughout the experiment. Control flasks were placed back in the incubator at ~100 $\mu E\ m^{-2}\ s^{-1}$. For all HL and control experiments, we used 3.7% HCl cleaned glass flasks while 3.7% HCl cleaned quartz flasks were used for all UV experiments. Stoppers made of cotton and gauze were used to close the flasks during all the experiments. All flasks were acid cleaned and autoclaved with the stoppers before use. The external water in the water bath reached to a level equivalent to about 700 ml of culture within the flask. Light levels, measured next to each flask and just above the

water line were between 460 and 590 μE m^{-2} s^{-1}. Experiments started at 10:30 (19.5 h experiments) and 12:30 (2.5 h experiments), in both cases 4.5 h after the incubator lights came on. FCM samples were taken in triplicate from each flask at 0 h, 2.5 h, 6 h, 9.5 h, and 19.5 h for the 19.5 h experiments. For the 2.5 h experiments, flasks were sampled for FCM at -4.5 h, 0 h, 1 h and 2.5 h, while RNA was harvested from 4 flasks at each of the following time points: 0 h, 1 h and 2.5 h.

Flow cytometry sampling and monitoring

During these experiments, cultures were monitored daily by FCM. One hunded μl of live culture was diluted into 890 μl of 0.2 μm filtered K artificial medium and 10 μl of yellow green 0.75 μm beads (Polysciences Inc., Warrington, PA, USA). Each sample was run on the flow cytometer (Epics XL, Beckman Coulter, Brea, CA, USA) for 1 min to fill the sample lines, then for 2 min for recording. Data acquisition was triggered on side scatter (SSC) red fluorescence and forward angle light scatter (FALS) and counts were also monitored.

During experiments, additional FCM samples were taken in triplicate from each flask for more precise analysis at a later date. One ml of culture was pipetted out of the flask, placed into a cryotube and 10 μl of glutaraldehyde (Tousimis, Rockville, MD, USA, final concentration 0.25%) was added to the tube. The tubes were quickly vortexed and placed in the dark at room temperature for 20 min. They were then cryo-frozen in liquid nitrogen for few hours to few days and transferred to -80 °C for long term storage. These samples were then run on a different flow cytometer (Influx, BD Biosciences, Franklin Lakes, NJ, USA). This instrument has greater FALS sensitivity for particles in the *Micromonas* size

range and samples can be run for a longer period allowing better statistical analysis of populations. Additionally, the instrument also has greater red fluorescence sensitivity and the beads we used as standards are not designed to fluoresce in the red (yellow-green beads, Polysciences Inc., Warrington, PA, USA), The Influx is equipped with a 488 nm laser (200 mW output), a 70 µm diameter nozzle and generally run at a flow rate of 25 µl min^{-1}. FALS, SSC (90° angle) and red autofluorescence from chlorophyll (692 ± 40 nm band-pass filter) were recorded for each sample.

RNA sampling

Flasks containing 1 l of culture were removed from the water bath. After FCM samples were taken from the flasks, the content was poured into a 1 l centrifuge bottle. All the centrifuge bottles were cleaned with RNase Zap solution (Ambion, Austin, TX, USA), thoroughly rinsed and dried before every use. At each time point, a total of 4 bottles (from 4 flasks) were centrifuged at 10,000 x *g* at 21 °C for 12 min (Sorvall RC26-Plus Superspeed, Thermo Scientific, Waltham, MA, USA). After the first centrifugation cycle, most of the supernatant (except for about 10 ml) was quickly removed (by pouring) from of each bottle. The pellet was resuspended in the remaining 10 ml of supernatant which were then pipetted into 1.7 ml RNase/DNase clean centrifuge tubes. The tubes were centrifuged at 10,000 x *g* at 21 °C for 10 min. All the supernatant was thoroughly removed with a pipet and the tubes were immediately placed on dry ice and transferred to -80 °C. The temperature of the culture after the spin was checked during test experiment and remained at <22 °C after the 12 min spin. The quality of cells in the pellet after the spin was monitored by observing scatter and fluorescence properties

which reflect the health of cells. In addition, an aliquot of the supernatant was also run on the flow cytometer and showed that about 6% of the cells remained in the supernatant after the spin.

RNA extraction and quantitative polymerase chain reaction (qPCR)

Samples were removed from storage at -80 °C and placed on dry ice until extraction. Total RNA was extracted using the Qiagen RNeasy kit (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions. Post-extraction, the RNA samples were treated with 1 µl of DNase at 37 °C for 30 min (TURBO DNA-*free*™ Kit, Ambion, Austin, TX, USA) according to the manufacturer's protocol in order to minimize genomic DNA contamination. Following DNase treatment, samples were purified by lithium chloride precipitation to remove additional DNA, proteins and carbohydrates. For this purification, a volume of 4.8 M LiCl (1 vol LiCl: 1 vol sample) was added and the sample was placed between 4 to 6 h at -20 °C. Following this step, samples were spun for 30 min at 4 °C and 16,100 x *g*. The sample was transferred to a new tube and 400 µl of 70% ethanol was added. After 10 min on ice, samples were spun for 5 min at 18,000 x *g* for 5 min. The last two steps were repeated, the supernatant was removed and the pellet was resuspended in TE buffer. All the samples were then analyzed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) pre- and post-treatment to determine RNA quality. The amount of total RNA was quantified on the Nano Drop system (Thermo Scientific, Waltham, MA, USA). Single stranded cDNA was made using the SuperScript® III First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions using oligo dT primers and 6 µl of total

RNA (12.5 ng μl^{-1}), for a total reaction volume of 60 µl. Negative control reactions without reverse transcriptase (RT- controls) were prepared for each sample. TaqMan primer and probe for the *Actin* gene, a commonly used endogenous control gene was designed by Applied Biosystems (Brea, CA, USA, McDonald et al., 2010) and for three target genes (*Lhcp1*, *Ohp2*, *Lhcsr*) were designed using Primer Expressed software 3.0 (Applied Biosystems, Brea, CA, USA, Table 4.1). All *Lhcp1*, *Ohp2*, *Lhcsr* probes were labeled with a fluorescent reporter FAM (6-carboxyfluorescein) at the 5′ end and a non-fluorescent quencher at the 3′ end (Table 4.1). The Actin probe was an Applied Biosystems Custom TaqMan® MGB™ Probe labeled with a fluorescent reporter FAM (6-carboxyfluorescein) at the 5′ end and a 3' non-fluorescent quencher (Table 4.1). This housekeeping gene (*Actin*) was chosen based on previous results from a selection of genes by McDonald et al. (2010) but was more stable than glyceraldehyde 3-phosphate dehydrogenase (GAPDH) chosen in by McDonald et al. (2010). The expression of *Actin* compared to the expression of GAPDH was assessed by qPCR and the difference in threshold cycle (C_T) between the various experimental conditions was determined. In this study, the *Actin* gene was chosen because it showed the least change in C_T between the different conditions.

All primers and probes set were compared to the entire *Micromonas sp.* RCC299 genome (using BLASTN) to confirm that only genes of interest were targeted. qPCR was performed using a 7500 Real Time PCR System (Applied Biosystems, Foster City, CA, USA) in MicroAmp Optical 96-well plates. Total reaction volume was 25 µl including 1 x Taqman Gene Expression Master Mix (Applied Biosystems, Foster City, CA, USA), 250 nM probe, 900 nM primers (final concentrations) and 2 µl (175 pg μl^{-1}) of cDNA.

Cycling parameters were 1 cycle of 50 °C for 2 min; 1 cycle of 95 °C for 10 min and 40 cycles of 95 °C for 15 sec followed by 60 °C for 1 min. Primer-probe sets were verified by running the qPCR product on 3% agarose gel with a 50 bp Mini ladder (Fisher Scientific, Pittsburgh, PA, USA) to verify fragment length. The linear dynamic range for each primers and probe set was tested using cDNA from samples from control experiment prepared from a serial dilution of RNA. The concentration of RNA added to the cDNA reaction fell within the linear part of the curve, equivalent to a 1:1 conversion between RNA and cDNA. The efficiency of primers and probe sets were determined using a dilution series of 1) qPCR product (purified with the MinElute PCR purification kit, Qiagen, Germantown, MD, USA) and 2) cDNA. C_T values were generated for all the treatment and control experiments and data were analyzed using the $2^{-\Delta\Delta C}{}_T$ method (Livak and Schmittgen, 2001) performed using the 7500 System SDS Software v1.4 (Applied Biosystems, Foster City, CA, USA), with *Actin* as the endogenous control and T_0 as the calibrator.

Table 4.1. Primer/probe set sequences used during this study. Protein ID refers to the ID in the *Micromonas sp*. RCC299 genome (v3.0, US Department of Energy's Joint Genome Institute, JGI, Worden et al. 2009). All *Lhcp1*, *Ohp2*, *Lhcsr* probes were labeled with a fluorescent reporter FAM (6-carboxyfluorescein) at the 5′ end and a non-fluorescent quencher at the 3′ end. The Actin probe was an Applied Biosystems Custom TaqMan® MGB™ Probe labeled with a fluorescent reporter FAM (6-carboxyfluorescein) at the 5′ end and a 3' non-fluorescent quencher.

Gene	Protein ID	Primer/Probe	Sequence 5'-3'
Lhcp1	113648	Forward	CGGAGCTTGAGTTGTCAGTTACTC
		Reverse	TCCAGCTTCGGCAAAACC
		Probe	CGGCGGTCGCTTTGACCCC
Lhcsr	105009	Forward	GCGACCACCGGCAACA
		Reverse	GACTTGACAGCCTCCTTGATGTC
		Probe	CAAGATCCAGCCCGGCAAGAAGTACG
Ohp2	106894	Forward	TCCTCGTGGGCATGATGAC
		Reverse	ACGGAGATGGTGAGCTTGATCT
		Probe	CCACCGGCGTGGACTTCATCG
Actin	90942	Forward	GCCCTCGTGTGCGATAAC
		Reverse	CCGACGATGGAGGGAAAGAC
		Probe	CCGGCCTTGACCATGC

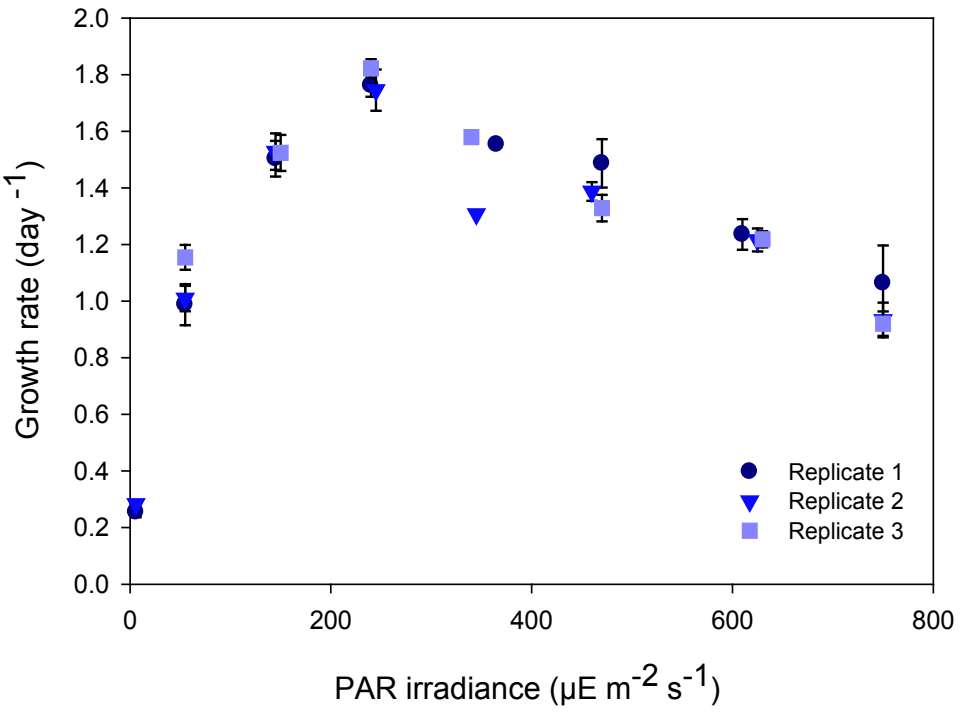
Results

Growth rates and cell physiology vs. PAR irradiance

Micromonas sp. strain RCC299 was grown axenically at different light levels to determine the effect of irradiance on growth rates (Figure 4.1). Cells were able to grow under the entire range of tested light levels, from 6 μE m^{-2} s^{-1} to 750 μE m^{-2} s^{-1}. The growth rate was lowest at 6 μE m^{-2} s^{-1} (on average: 0.27 ± 0.02 day^{-1}) and μ_{max} (on

average: 1.78 ± 0.05 day^{-1}) was reached at 240 µE m^{-2} s^{-1}. Beyond that light level, growth rates decreased but remain relatively high, even at 750 µE m^{-2} s^{-1} (0.97 ± 0.08 day^{-1}).

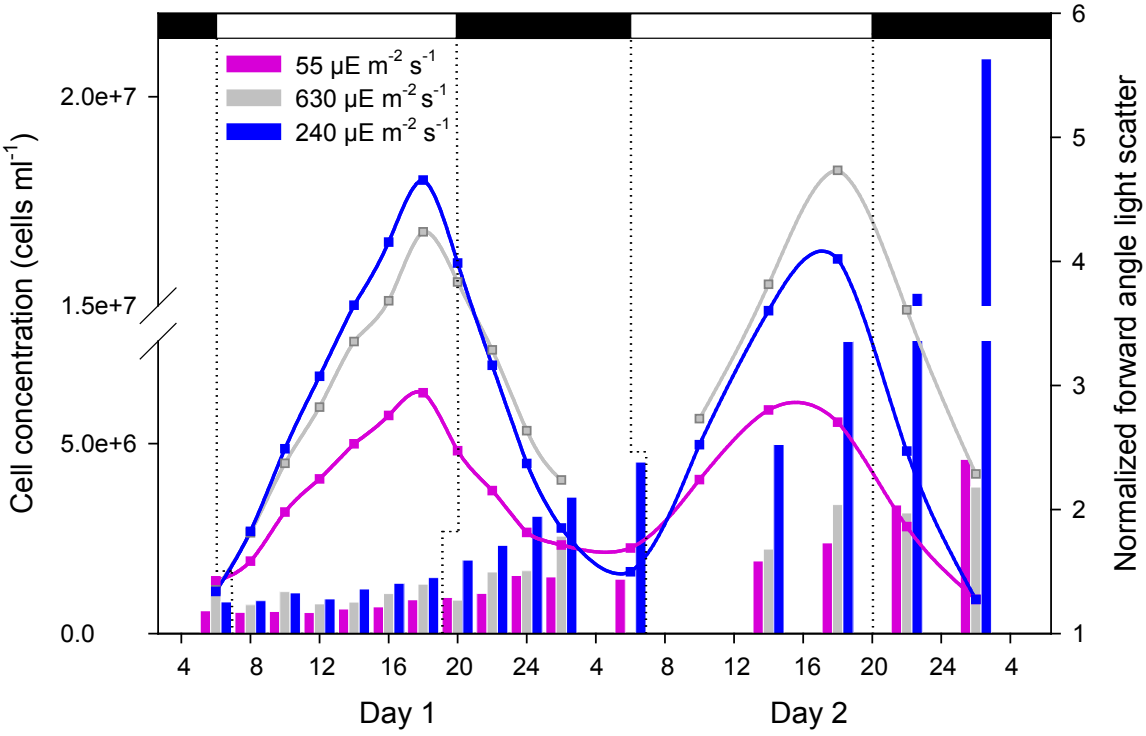
Figure 4.1. *Micromonas sp.* RCC299 growth rate (day^{-1}) vs. photosynthetic active radiation (PAR) irradiance (µE m^{-2} s^{-1}). Axenic cultures were grown in K medium made with artificial seawater and acclimated to each light level for at least 10 generations. Triplicate cultures were grown at each light level. Average growth rates of two to 12 transfers (depending on light level) after 10 generations were calculated for each culture. Bars represent standard error of transfers.



Results show clear variations in cellular characteristics over the course of the 50 h. This is expected given synchronization of these photosynthetic cells to the L:D period. Results show apparent variations in cellular characteristics over the course of the 50 h.

FALS (normalized to bead FALS) increased consistently throughout the light period from 6 h to 18 h and decreased during the dark period from 18 h to 6 h, to a minimum value at 6 h (Figure 4.1). Note, the data presented does not reflect transfer close to the mid-point, as cell concentrations were adjusted for the volume of media added. The peak in FALS occurred just before light: dark transition on both days (Figure 4.2). This trend was observed for all three light levels and for both sampling days. Normalized FALS was lowest for the cells grown at 55 $\mu E\ m^{-2}\ s^{-1}$ (the lowest light level). Cells at this light level also grew more slowly that cell at higher light levels (Figure 4.2, bars). Cultures were able to grow at all light levels, as shown by the increase in cell concentrations throughout the 50 h of sampling. Growth rates on day 1 and day 2 were compared to growth rates obtained previous to the experiment (Figure 4.1) to confirm that the sampling did not affect cells. Growth rates were around 1.10 day^{-1}, 1.76 day^{-1} and 0.72 day^{-1} for cultures under 55, 240 and 630 $\mu E\ m^{-2}\ s^{-1}$ of irradiance respectively. These numbers fell within expected ranges for these light levels except for 630 $\mu E\ m^{-2}\ s^{-1}$. Average growth rates after 10 generations but prior to sampling were around 1.15 ± 0.10 day^{-1}, 1.82 ± 0.06 day^{-1} and 1.22 ± 0.09 day^{-1} for 55, 240 and 630 $\mu E\ m^{-2}\ s^{-1}$ respectively (Figure 4.1). The discrepancy between rates reported prior and during the experiment for 630 $\mu E\ m^{-2}\ s^{-1}$ was due to the significantly variation in rates between day 1 (1.12 day^{-1}) and day 2 (0.41 day^{-1}), indicating that cells were no longer in mid-exponential phase on day 2. Cell numbers increased most during the night and growth rates were higher during the dark period than during the light period for all light levels.

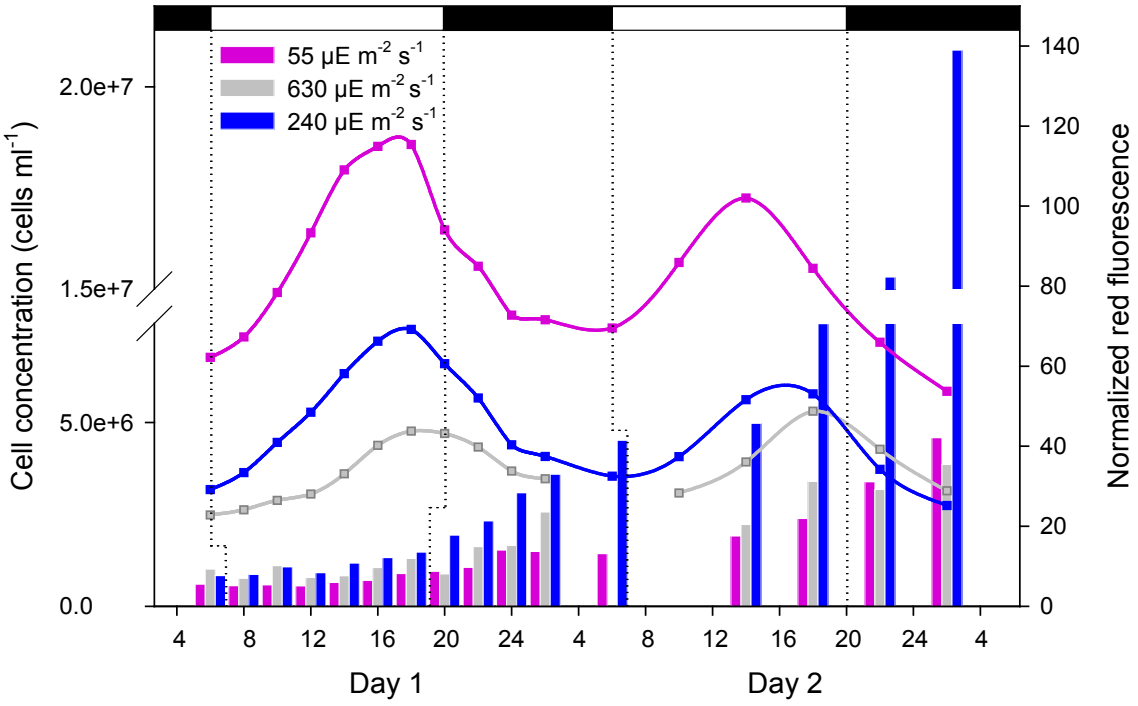
Figure 4.2. Cell concentration (bars, cell ml^{-1}) and FALS (normalized to bead FALS, lines, relative units) for *Micromonas sp.* RCC299 grown axenically at three light levels: 55 µE m^{-2} s^{-1}, 240 µE m^{-2} s^{-1} and 630 µE m^{-2} s^{-1} over a period of 48 h. Samples were taken every 2 h for 0-24 h (6:00 to 6:00) and every 4 h for 24-44 h (6:00 to 2:00). Cultures were first acclimated to each light level for at least 10 generations before being sampled. In addition, cells were transferred to fresh media (i.e. diluted with fresh media) between 18 and 20 h on day 1 to keep them in exponential phase. Concentrations of cells represented on the graph have been adjusted so that the dilution is not reflected. Cultures were grown on a 14:10 L:D cycle, denoted by the black (night) and white (day) bars (top). Red fluorescence of cells was normalized to red fluorescence of 0.75 µm yellow-green beads. Cell concentrations at 38 h were not determined.



The daily patterns of normalized red fluorescence (representing mean chlorophyll fluorescence per cell) followed a diel cycle (Figure 4.3), similar to the FALS patterns (Figure 4.2). A constant increase in red fluorescence corresponding to chlorophyll synthesis inside the cell was observed during the subjective day period followed by a

regular decrease during the night period. On day 2 of the experiment, maximum red fluorescence occurred earlier during the day than on day 1, peaking between 14 h and 18 h, compared to a peak around 18 h on day 1. Lowest levels of red fluorescence for each culture were always observed at 6 h. Overall, the culture growing in the lowest light level had higher normalized red fluorescence. This suggests that cells increased their amount of chlorophyll as the amount of light available decreased.

Figure 4.3. Cell concentration (bars, cell ml^{-1}) and red (chlorophyll) fluorescence (normalized to bead red fluorescence, lines, relative units) for *Micromonas sp.* RCC299 grown axenically at three light levels: 55 µE m^{-2} s^{-1}, 240 µE m^{-2} s^{-1} and 630 µE m^{-2} s^{-1} over a period of 48 h. Samples were taken every 2 h for 0-24 h (6:00 to 6:00) and every 4 h for 24-44 h (6:00 to 2:00). Cultures were first acclimated to each light level for at least 10 generations before being sampled. In addition, cells were transferred to fresh media (i.e. diluted with fresh media) between 18 and 20 h to keep them in exponential phase. Concentrations of cells represented on the graph have been adjusted so that the dilution is not reflected. Cultures were grown on a 14:10 L:D cycle, denoted by the black (night) and white (day) bars (top). Red fluorescence of cells was normalized to red fluorescence of 0.75 µm yellow-green beads. Cell concentrations at 38 h are not determined.

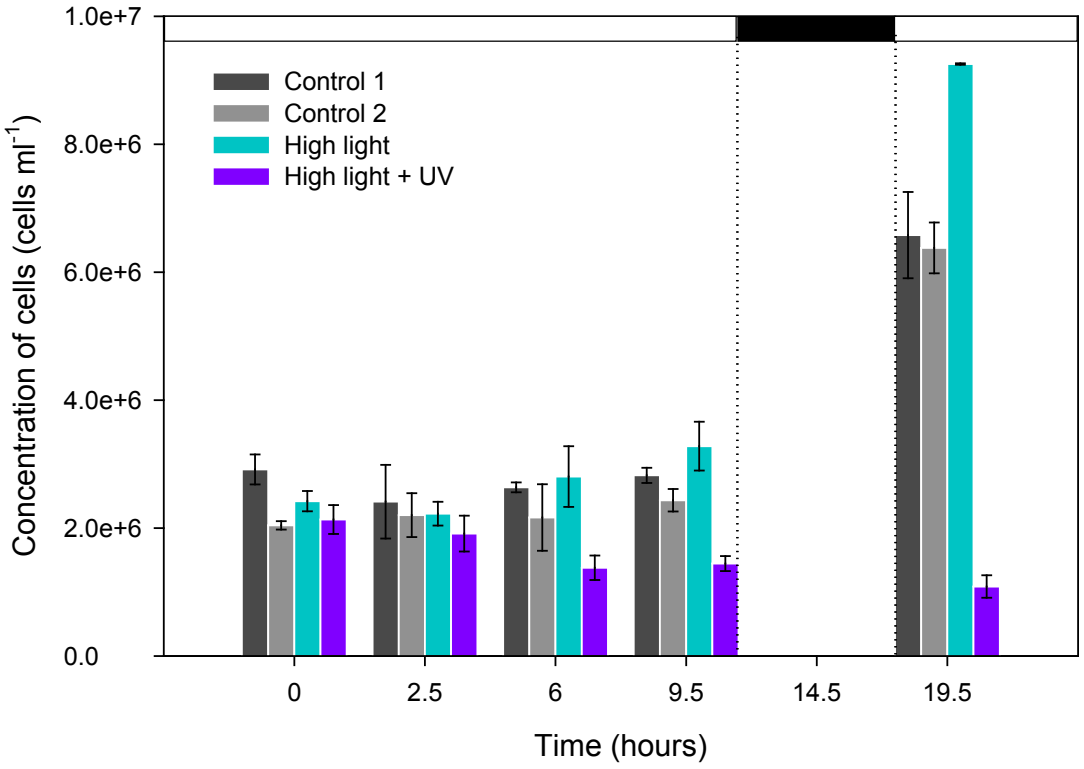


Effect of high light and UV radiation on growth and cellular parameters

Cultures grown at 100 µE m^{-2} s^{-1} for at least 10 generations were subject to a light shift at t=0 h (4.5 h after the beginning of the subjective day). Cells were exposed to two different treatments (HL or HL + UV) for 19.5 h. Cellular parameters as well as cell

concentration were monitored throughout the experiment. Concentration of cells for the control experiments were constant during the light period but had increased significantly by the end of the subjective night. Growth rates over the course of the experiments were between 1.00 day^{-1} and 1.40 $day^{-1}$for control cultures. Cells shifted to HL only grew faster (1.65 day^{-1}), while concentration of cells subjected to a HL + UV treatment decrease constantly throughout the experiment (-0.83 day^{-1}, Figure 4.4).

Figure 4.4. Concentration of *Micromonas sp.* RCC299 cells (cells ml^{-1}) in control, HL and HL + UV treatment experiments over 19.5 h. Cells were grown at 250 µE m^{-2} s^{-1} for several weeks (and at least 10 generations). On the day of the experiments, flasks were placed in their respective treatment at t=0 h (4.5 h after the light came on). Cultures were grown on a 14:10 L:D cycle and this regime, noted by the black (night) and white (day) bars (top), was kept during the experimental treatment. Data represent the average multiple flasks and error bars represent standard deviation of the biological replicates.



Percent change in normalized FALS and normalized red fluorescence compared to t=0 h were also recorded. Control experiments showed a pattern similar to cell cycle observed in Figure 4.2 and 4.3. FALS (Figure 4.5) and red fluorescence (Figure 4.6) increased consistently during the day period to a maximum (152% and 120% increase for FALS and 60% and 71% increase for red fluorescence for control 1 and control 2, respectively) reached just before the L:D transition. At t=19.5 h (6:00), just after the night period, FALS and red fluorescence was at a minimum, i.e. relatively lower than at the beginning of the experiment (10:30). In the HL treatment, normalized FALS followed the same trend as rather than in the control, with a maximum FALS (155% increase) just before the subjective night. In the HL treatment, cells appeared to start readjusting their chlorophyll content as early as 2.5 h after treatment exposure. Percent change in red fluorescence at all time points were significantly different that of the control. Even though red fluorescence increased (or stay constant the first 9.5 h of the experiment) cells modified their chlorophyll content throughout the whole experiment. In the HL + UV treatment, red fluorescence was significantly lower than the control at all time points as well. However, the chlorophyll content decreased rapidly to a concentration lower than the start of the experiment.

Figure 4.5. Percent change in normalized FALS in HL, HL + UV and two control experiments over 19.5 h. Cells were grown at 100 µE m^{-2} s^{-1} for several weeks (and at least 10 generations). On the day of the experiment, flasks were placed in their respective treatment at t=0 h. Cultures were grown on a 14:10 L:D cycle and this regime was kept during the experimental treatment. * denote samples that were significant different from control (HL vs. control 1, HL + UV vs. control 2, control 1 vs. control 2). Data represent the average multiple flasks and error bars represent standard deviations of biological replicates.

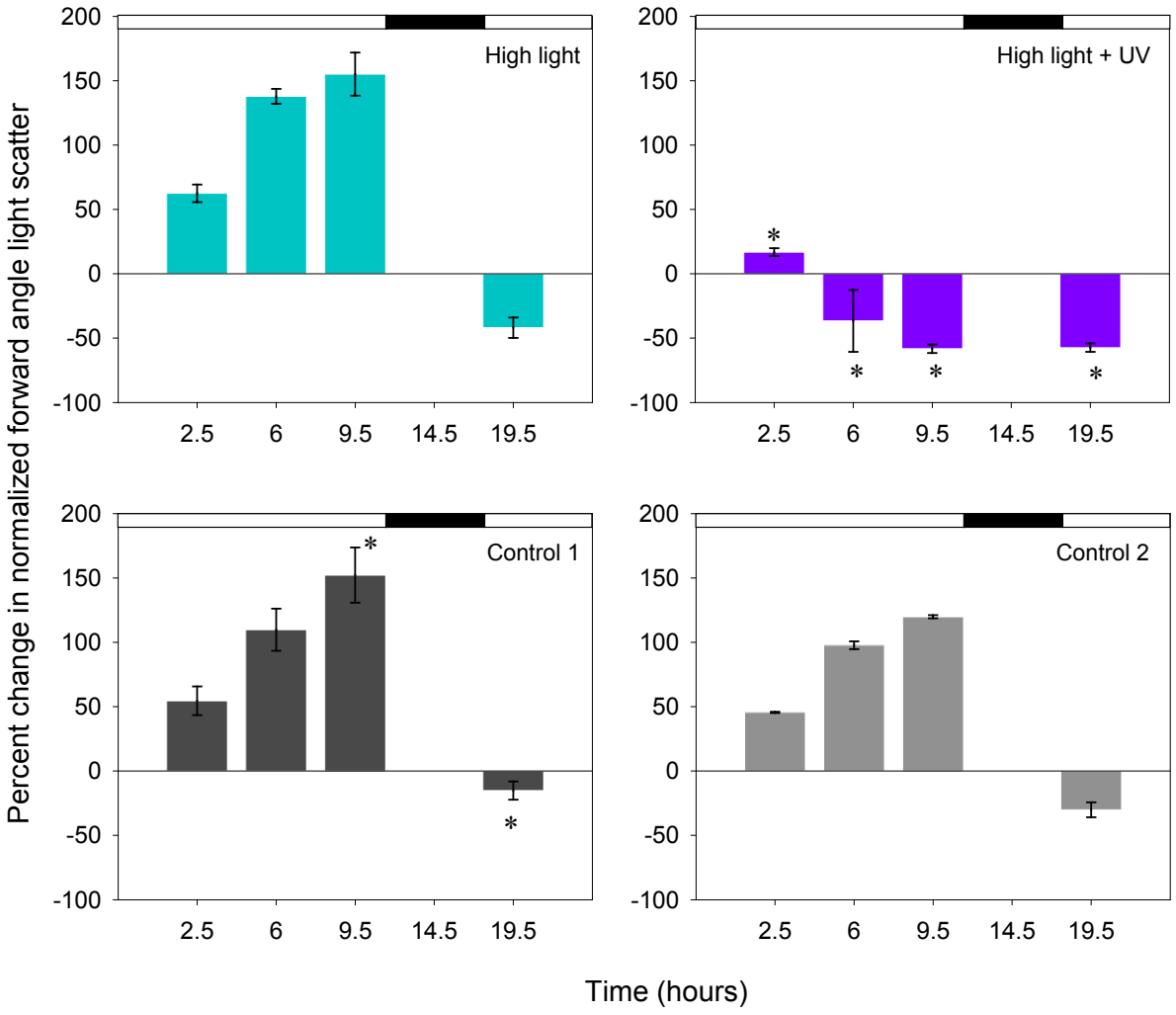
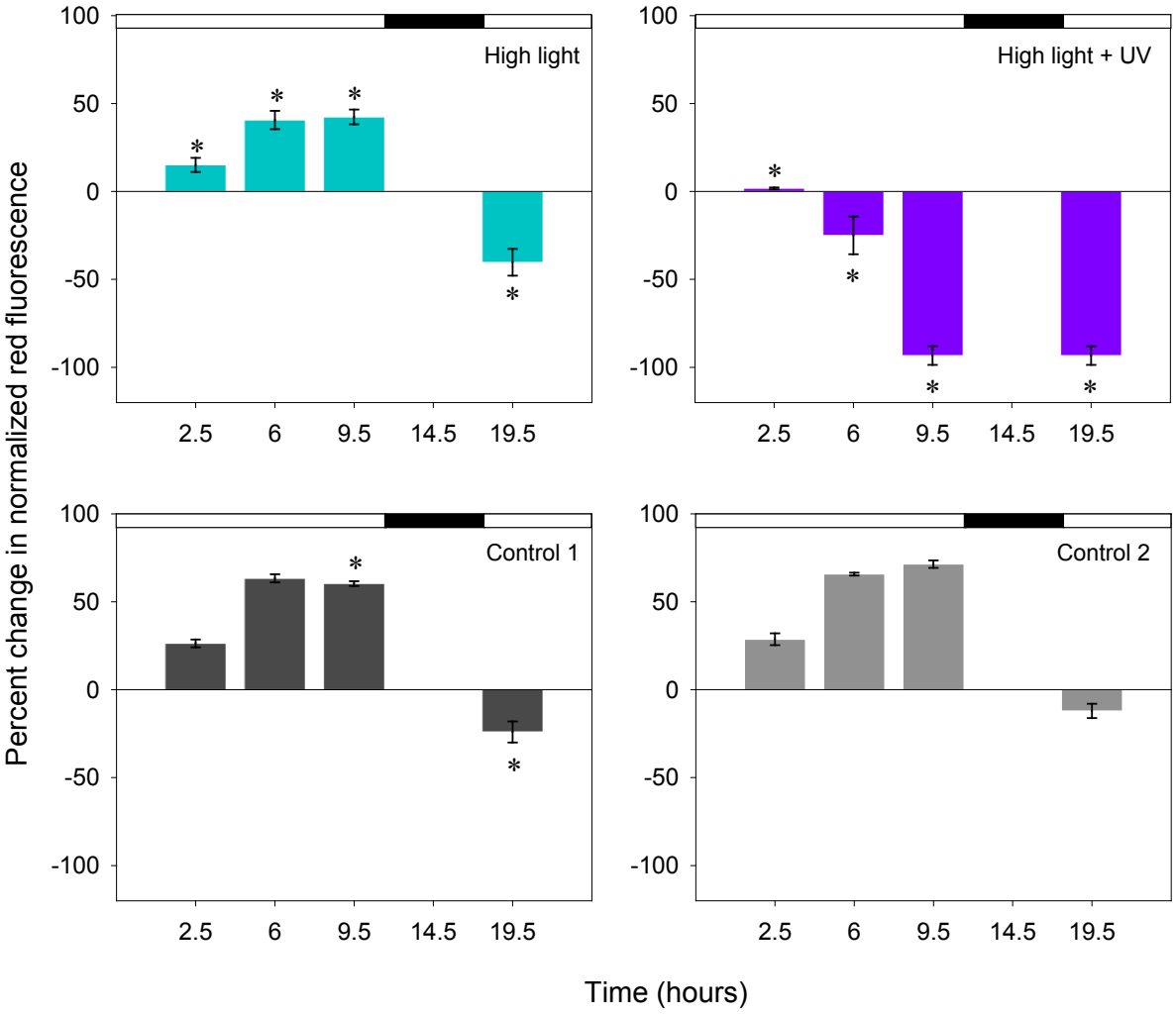
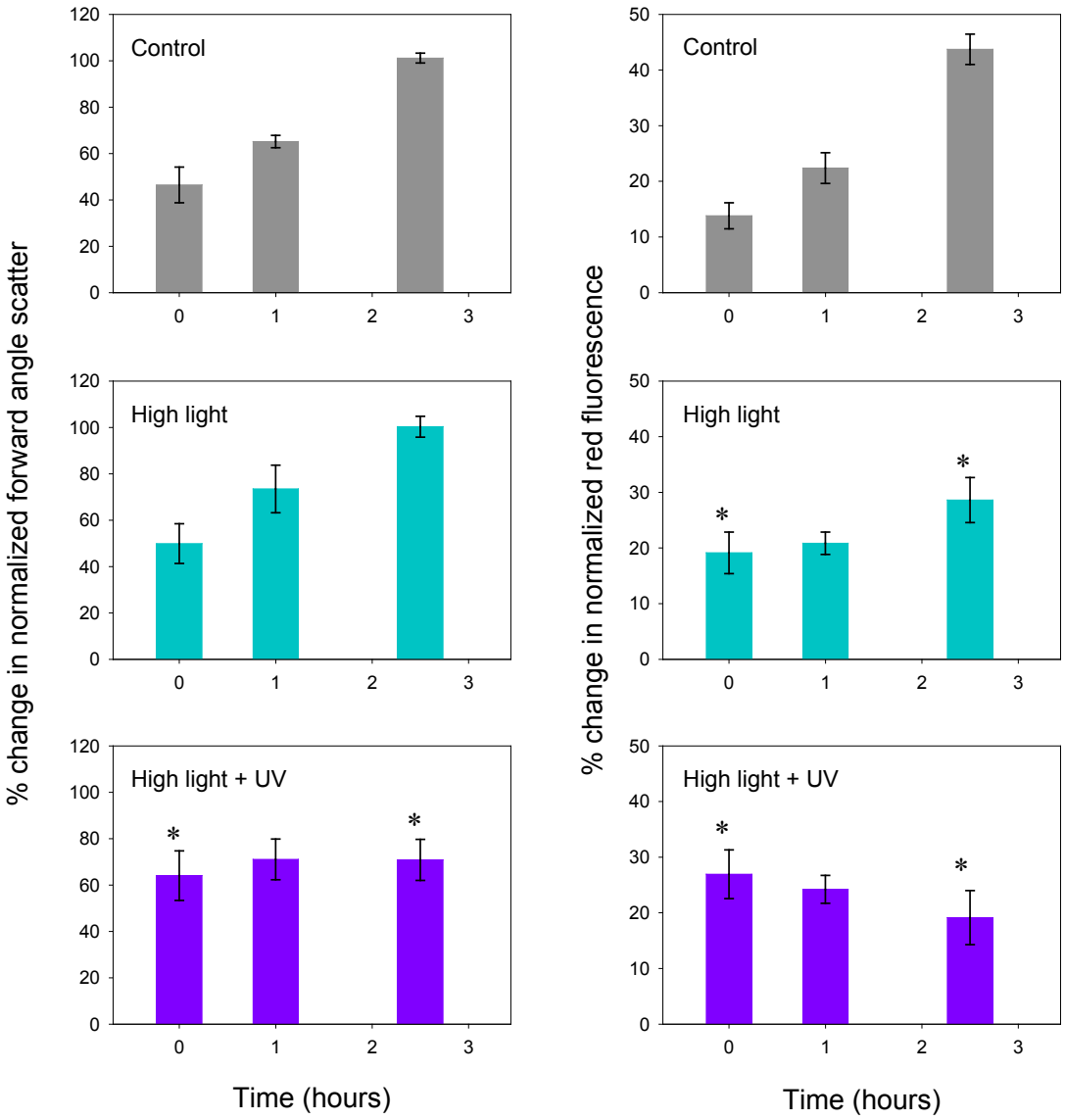


Figure 4.6. Percent change in normalized red (chlorophyll) fluorescence in HL, HL + UV and two control experiments over 19.5 h. Cells were grown at 100 µE m^{-2} s^{-1} for several weeks (and at least 10 generations). On the day of the experiment, flasks were placed in their respective treatment at t=0 h. Cultures were grown on a 14:10 L:D cycle and this regime was kept during the experimental treatment. * denote samples that were significant different from control (HL vs. control 1, HL + UV vs. control 2, control 1 vs. control 2). Data represent the average of multiple flasks and error bars represent standard deviations of biological replicates.



Trends in cell physiology in the shorter experiments (2.5 h) were similar to that in the 19.5 h experiments. In the control, normalized FALS and red fluorescence increased at each time point (Figure 4.7). Percent change in normalized FALS was not significantly different in the HL and in the control. In the HL + UV treatment, cells FALS slightly increased in the first hour but did not follow the trend of healthy cells and by 2.5 h, normalized FALS was significantly different than the control. Red fluorescence in the HL treatment was higher at t=2.5 h than at t=0 h, however, cells had already started to readjusted their chlorophyll content since there was a statistical difference with the control. In the HL + UV treatment the chlorophyll content per cell decreased constantly and was significantly different than the control as well (Figure 4.7).

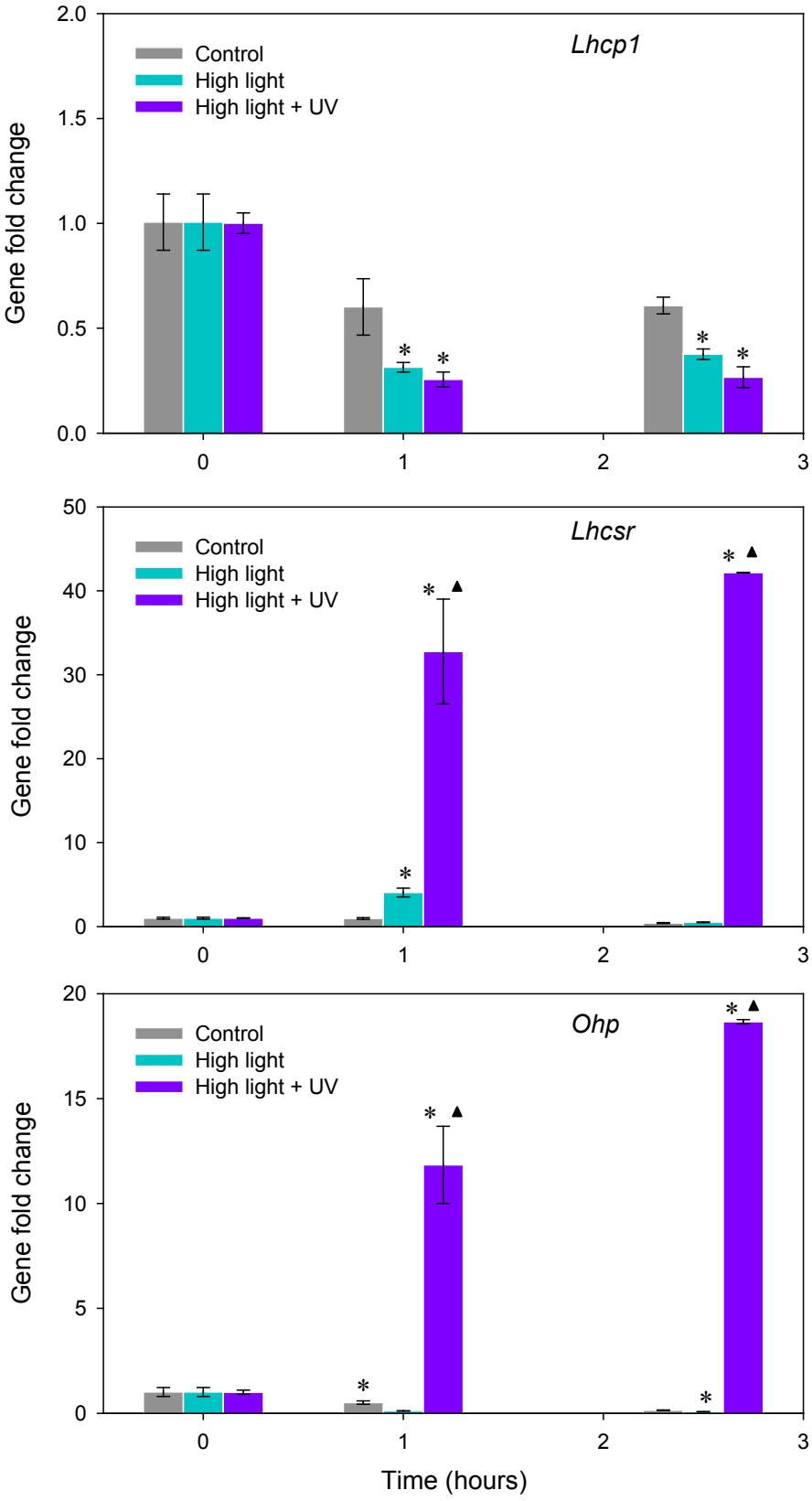
Figure 4.7. Percent change in normalized FALS (left) and normalized red (chlorophyll) fluorescence (right) for three experiments: control (top), HL (middle) and HL + UV (bottom). Data represent the percent change relative to t=-4.5 h (the time the lights came on). Cells were grown at 100 µE m^{-2} s^{-1} for several weeks (and at least 10 generations). On the day of the experiment, samples were taken at t=-4.5 h. Cells were then transferred to flasks and place in their respective treatments at t=0h. Data represent the average of multiple flasks and error bars represent standard deviations of biological replicates.



Gene expression

During the 2.5 h experiments, we also monitored the expression of three genes. One of those genes (*Lhcp*) encodes for light harvesting complex proteins thought to be specific to prasinophytes. The second gene *Lhcsr* encodes for proteins critical for the survival of the green algae *C. reinhardtii* in a dynamic light environment (Peers et al., 2009). The third gene, *Ohp2,* is thought to be involved in photoprotection. Data analyzed by the $2^{-\Delta\Delta C}{}_T$ method showed that *Lhcp* expression decreased in all the experiments. In the control, expression at t=1 h decreased 0.6 fold from what it was at t=0 h and remained constant at the next time point. In the HL treatment, expression was significantly lower than in the control with expression around 0.3 to 0.4 fold compared to the beginning of the experiment. Expression of *Lhcsr* in the HL experiment reached a maximum at t=1 h (4.0 fold higher than at t=0 h) and was slightly lower at t=2.5h than at starting expression. The expression of this gene was much higher in the HL + UV treatment, where it increased to 33 fold its initial expression at t=1 h and to 42 fold at t=2.5 h. The expression of *Ohp2* in the HL + UV increased consistently as well over time and reached a maximum at t=2.5 h (19 fold increased compared to t=0 h). Interestingly this trend was not observed in the HL treatment, where the expression of *Ohp2* decreased over time, to levels significantly lower than in the control.

Figure 4.8. Gene fold change in two treatments (HL and HL + UV) and control experiments. C_T values were generated for all the treatment and control experiment and data were analyzed using the $2^{-\Delta\Delta C}{}_T$ method (Livak and Schmittgen, 2001) with *Actin* as the endogenous control and T_0 as the calibrator. * denote samples that were significantly different from control, data for *Lhcsr* and *Ohp2* were log transformed. ▲ denote samples that were significantly different from high light. Data represent the average of multiple flasks and error bar represent standard deviations of biological replicates.


Lhcp1
Control
High light
High light + UV
Gene fold change
2.0
1.5
1.0
0.5
0.0
0
1
2
3
Lhcsr
Control
High light
High light + UV
Gene fold change
50
40
30
20
10
0
0
1
2
3
Ohp
Control
High light
High light + UV
Gene fold change
20
15
10
5
0
0
1
2
3
Time (hours)

Discussion

Growth vs. Irradiance

Micromonas cells were able to grow at light levels from 5 μE m^{-2} s^{-1} to 730 μE m^{-2} s^{-1}, with μ_{max} attained at ~250 μE m^{-2} s^{-1} (Figure 4.1). These results suggest that *Micromonas sp.* RCC299 is able to colonized environments with a wide distribution of light levels. *Ostreococcus* strain OTH95 and RCC501 were able to grow at similar irradiances, with μ_{max} slightly lower than *Micromonas sp.* RCC299 (Rodríguez et al., 2005), although media and seawater differences may affect the validity of direct comparison of our results to those of Rodríguez et al. (2005). Our results suggest that this strain of *Micromonas* is optimized for faster growth at irradiance 50 μE m^{-2} s^{-1} or higher (Figure 4.1). The highest light levels tested in this study did not severely compromise the growth of the respective cultures, even though growth rates recorded at irradiances greater than 250 μE m^{-2} s^{-1} continuously decreased. Rodríguez et al. (2005) suggested that HL and low light ecotypes of *Ostreococcus* isolates exist (see also Cardol et al., 2008), strains OTH95 and RCC501 belonging to the HL type. Whether ecotypes of *Micromonas* that are differently light adapted exist is currently unknown. Multiple strains would need to be compared and their PS characteristics and capacities analyzed to determine whether light plays a role in ecotypic differentiation for this genus.

Cell cycle

Synchronization of cell cycle during growth under a L:D cycle has been observed in green algae (e.g. unicellular eukaryotic green alga *C. reinhardtii*, Bruce et al., 1970) and picoeukaryotes, including *Ostreococcus* (Farinas et al., 2006) and *Micromonas sp*.

CCMP490 (formerly RCC114, Jacquet et al., 2001). In this study, we also observed synchronization of the cell cycle under growth conditions similar to natural environment (i.e. L:D cycle). During the light phase, average chlorophyll fluorescence per cell increased (Figure 4.3), consistent with the synthesis of chlorophyll in the growing chloroplast (Moulager et al., 2007). In *Ostreococcus*, a decrease of red fluorescence per cell is detected upon division of the two chloroplasts, and is directly correlated to cell division. When cells are in stationary phase and do not divide, the decrease in the chlorophyll fluorescence at the end of the light period is not detected (Moulager et al., 2007). In our study, a decrease in chlorophyll fluorescence was observed at the end of the light period for cultures grown at three different light levels (Figure 4.3), indicating that cells start to divide at the beginning of the dark period. Similar results were observed for *Micromonas* by Jacquet et al. (2001). In addition, cell division occurs when cells reach a critical size, presumably facilitated by the accumulating energy resulting from light-period photosynthesis (Moulager et al., 2007). Observations of normalized FALS confirmed that *Micromonas* increased in cells size and cell carbon (DuRand and Olson, 1998) during the light period (Figure 4.2). The decrease of normalized FALS starting at the end of the light period and throughout the dark period is consistent with a cell division occurring at the beginning night (Jacquet et al., 2001). The synchronization of cell cycle was observed for cultures acclimated for multiple generations at three different light levels (Figure 4.2, 4.3). However, cells grown at the lowest light levels (55 $\mu E\ m^{-2}\ s^{-1}$) consistently had high chlorophyll fluorescence indicating a higher amount of chlorophyll per cell (Figure 4.3). This strike out that more pigments are involved in the

light harvesting and antennae size is (Niyogi, 1999), as expected for growth under reduced light conditions. In contrast, cells grown at 55 μE m^{-2} s^{-1} had the lowest normalized FALS (Figure 4.2) and therefore lower cellular carbon content.

Effect of high light and UV radiation on growth and cellular parameters

Cells characteristics and growth rates of cells shifted to HL and HL + UV were observed in multiple experiments. Results showed that percent change in normalized FALS were similar in control and HL experiments in the 2.5 h and 19.5 h experiments (Figure 4.5, 4.7). This indicates that *Micromonas* cell size does not vary significantly when exposed to HL. This is in contrast with cells acclimated for at least 10 generations to different light levels, as normalized FALS of cells grown at the lowest irradiance (55 μE m^{-2} s^{-1}) was considerably less than for higher light levels. However, in the HL + UV treatment, FALS decreased through the 19.5 h, observed as soon as after 1 h of UV treatment (Figure 4.5, 4.7). Cell size therefore appears to be significantly affected by UV radiation but not HL, unless combinatorial effects were different than each independently. Cellular red fluorescence was affected by both the shift to HL and the shift to HL + UV. However, red fluorescence in the HL treatment continued to increase, consistent with cellular growth (but significantly less than the control) while red fluorescence in the UV decreased right away (Figure 4.7). In the green alga *Chlamydomonas*, chlorophyll monitored by FCM showed a more dramatic change than in *Micromonas*, as chlorophyll per cell decreased by ~20% during the first 6 h of HL (500 μE m^{-2} s^{-1}) exposure (Baroli et al., 2004) and by 50% within 24 h of exposure to 750 μE m^{-2} s^{-1} (Durnford et al., 2003). This implies that the number of photosynthetic complexes

after a shift to HL may be more reduced in *Chlamydomonas* than in *Micromonas* (Durnford et al., 2003). Harris et al. (2009) showed that in the prymnesiophyte *Emiliania huxleyi*, net chlorophyll *a* synthesis stopped after a shift from low light to HL. Nevertheless, cell division continued for 36 h after initiation of the experiment. The authors therefore suggest that cellular chlorophyll, after a shift to HL, was "diluted' by cell division rather than degraded. Similar results were observed in the marine diatom *Thalassiosira weisflogii* and *Skeletonema costatum* where chlorophyll was not rapidly degraded but instead rate of synthesis decrease and cell division were responsible for the lower pigment content per cell (Post et al., 1984, Anning et al., 2000).

Gene expression

One of the ways algae and cyanobacteria respond to light stress is by producing photoprotective proteins. Here, we compared the expression of a light harvesting complex gene specific for prasinophytes, *Lhcp1*, and two stress related genes: *Lhcsr* and *Ohp2* under various treatments. LHCP are nuclear-encoded proteins likely to have functions similar to the major PSII antennae in other green algae (Six, Worden et al., 2005). The expression of *Lhcp1* was monitored during a shift to HL. Like the control, but to a greater extent, a reduction in *Lhcp1* transcripts was observed presumably caused either by the interruption in transcription or by degradation. Within the first hour of shift to HL, as well as to the shift to HL + UV, the level of gene expression fell to 20-30% of the initial level (Figure 4.8). After 2.5 h of HL or HL + UV exposure, *Lhcp1* transcripts slightly increased compared to 1 h. Similar results were observed in *Chlamydomonas*; mRNA abundance of *LhcII-4* and *Lhcb4*, two genes encoding major and minor proteins

of the PSII respectively, decreased by 20-30% within 2 h of a shift to HL (Teramoto et al., 2004). In another study on *Chlamydomonas Lhcb* (LHCII) mRNA abundance dropped by 65% within 2 h following the shift to HL but recovered to low light levels within 8 h (Durnford et al., 2003). Our results are also consistent with the decrease of *Lhc* transcript observed in *Dunaliella salina* (Masuda et al. 2003). These rapid acclimation responses to HL 'stress' suggest a decreasing antenna size (Teramoto et al., 2004). In *Chlamydomonas*, the *Lhc* genes expression depend on the light energy absorbed by LHC and the energy utilized by CO_2 assimilation (Teramoto et al. 2002).

As mentioned above, in the control, a smaller change in gene expression was observed, but *Lhcp1* transcript still decreased within the first hour of the experiments (Figure 4.8). Circadian control of gene expression was first discovered in plants (Kloppstech, 1997). Here, the results for the genes that were monitored also show some diel oscillations, even within a 2.5 h time scale, as other *Lhc* and photosynthetic genes have been observed to do the same. For example in the cyanobacteria *Prochloroccocus*, grown in a L:D cycle, transcript of the *psbA* (a chloroplast encoded gene) was strongly correlated with the light, and peaked in the middle of the light period (Garczarek et al., 2001). Monnier et al. (2010) showed that in *Ostreococcus*, 80% of expressed genes had rhythmic patterns of expression and genes related to photosynthesis varied cyclically. In *Chlamydomonas*, 9 *Lhc* genes possibly had circadian control (Kucho et al., 2005).

The expression of *Lhcsr* and *Ohp2* were also evaluated after the experimental shifts to HL or HL + UV. LHCSR was recently shown to be critical for the survival of *Chlamydomonas* under shifting light conditions. OPH proteins are thought to be involved in photoprotection and energy dissipation as opposed to light harvesting, although their

exact roles and functions remained to be determined. Here, *Lhcsr* expression studied by qPCR revealed that this gene was highly expressed under HL or HL + UV stress. One hour after the shift to treatment conditions, the amount of transcript was already 4 and 33 fold higher in HL and HL + UV, respectively, compared to the initial amount (Figure 4.8). In HL conditions, mRNA *Lhcsr* decreased to low levels after two hours, while within the HL + UV, transcription appears to continue to increase. These results suggest that in *Micromonas*, similar mechanisms may be responsible for triggering this gene, in both HL and HL + UV conditions. In *Chlamydomonas*, a *Lhcsr* gene (*LI818r-1*) was expressed at high levels and increased within 30 min of transfer to HL. The level of transcript then declined within 1 to 3 h of exposure to HL (Leford et al., 2004). We observed similar results with the decline of *Lhcsr* transcript after 2.5 h. Furthermore, *Lhcsr* mRNA and LHCSR protein were higher in *Chlamydomonas* cells grown in HL and an *Lhcsr* deficient mutant were unable to dissipate excess light energy by de-excitation of chlorophyll molecules in the PSII (Peers et al., 2009).

We also analyzed the expression of *Ohp2*, one of the 2 *Ohp* genes present in *Micromonas sp.* RCC299. The expression of *Ohp2* was similar to *Lhcsr* in the HL + UV treatment, and mRNA levels continued to increase from 0 to 2.5 h. Interestingly, *Ohp2* was not induced by HL treatment on its own (Figure 4.8). Results showed that the expression of *Ohp2* in HL was 'opposite' to the expression in HL + UV, decreasing significantly over time. This indicates that in *Micromonas* the induction of *Ohp2* is triggered by physiological impacts associated with the UV treatment but not by the HL treatment. On the basis of the quantitative data presented here, we can conclude that UV triggers a different pathway or that the cells are relatively resistant to high levels of light.

It should be noted that since we did not observe severe photoinhibition or bleaching in our growth versus irradiance experiments, it may be that with higher light levels than achieved here, that did have such consequences, cellular responses would be different. Our results are at odds with those of Andersson et al., 2003 who demonstrated that in the plant *Arabidopsis* OHP2 proteins present in the thylakoid membranes under low-light conditions increased after a shift to HL. Perhaps more directly comparable to our data, which was based on transcript analysis, the accumulation of *Ohp2* transcript (and protein) was only induced under HL; other stress conditions such as UV-A, heat, cold, desiccation or oxidative stress did not up-regulate the expression of this gene (Andersson et al., 2003). Other studies also reported the up-regulation of *Ohp1* gene under HL in *Arabidopsis* (Jansson et al., 2000). *Lhl2* (*Ohp1*) in *Chlamydomonas* was rapidly induced after a shift from LL to HL within 1 h and levels decreased back to initial level within 6 h (Teramoto, et al. 2004). In the cyanobacterium *Synechocystis*, various *hlip* transcripts, encoding for OHP in cyanobacteria increased under HL and UV-B (Dolganov et al., 1995, He et al., 2001, Huang et al, 2002). Nevertheless, Huang et al, 2002, showed that *hliA*, *hliB* and *hliC* were all highly induced by UV-B, and while *hliC* was induced under light intensities of 200 μE m^{-2} s^{-1}, the induction of *hliA* and *hliB* required a light level of 500 μE m^{-2} s^{-1}. It is therefore possible that the 'high' light level used during our experiments was not high enough, as discussed above, to trigger the expression of *Ohp2*. Alternatively, our results indicate that the induction of *Ohp2* gene requires UV radiation, or is more essential to photoprotection under UV conditions, which probably activates specific receptors. Kimura et al. (2003) confirmed that *Ohp* did not respond to HL conditions in *Arabidopsis*, however, down-regulation was not observed either.

In summary, this research establishes that two genes putatively involved in photoprotection or photoacclimation were induced differently in our experimental conditions, while a more classical light harvesting gene rendered different responses. Whether the difference in gene expression reflects actual difference in protein levels, or function of the antennae system, remains to be determined by future studies.

Chapter 5:

Conclusion

This dissertation research made a number of contributions to our understanding of phytoplankton and especially to the field of picoeukaryotic diversity and ecology. The various approaches used ranged from molecular tools such as polymerase chain reaction (PCR), fluorescence *in situ* hybridization (FISH), 18S rRNA gene clone libraries, to techniques including flow cytometry, microcopy and dilution experiments. These multiple approaches, in combination with several collaborations, have contributed to a better appreciation of picoeukaryotic contributions in marine environments.

Overall, this research emphasizes and confirms that picoeukaryotes (eukaryotes $\leq$ 2-3 μm in diameter) are a significant and consistent component of the picophytoplankton in marine environments. In addition to *Prochlorococcus* and *Synechococcus*, the other two dominant groups of the picophytoplankton, phototrophic picoeukaryotes are widespread in various ocean basins. In chapter 3, picophytoeukaryotes were shown to be present in the Sargasso Sea where they formed a substantial portion of the biomass. This was also the case in the Florida Straits (chapter 2 and 3) were picoeukaryotes were present year round with some variability in abundance. As often observed in oligrotrophic waters (e.g. Blanchot et al., 2001), picoeukaryotes were less abundant than *Prochlorococcus* and *Synechococcus* at most of the sites and depths investigated herein.

Our results also strengthen data supporting the idea that picoeukaryotes are indeed very diverse and that picoeukaryotic diversity is still being revealed. In the last few years, clone libraries from environmental samples have allowed us to highlight diversity that

had largely been overlooked until then (e.g. Massana and Pedós-Alió, 2008, Worden and Not, 2008). The use of molecular tools emphasized the fact that traditional methods simply missed a lot of this diversity, as new taxa of picoeukaryotes with no cultured representative have been discovered. This is especially true for photosynthetic picoplankton because clone libraries are thought to be heavily biased towards heterotrophic cells (Vaulot et al., 2008). Our data built upon the idea that many uncultured organisms have been unnoticed and have yet to be isolated and identified.

In the first part of the dissertation, the discovery of biliphytes at Bermuda Atlantic Time-series Study (BATS), a well-studied time-series site, reflects the current paucity of our knowledge concerning eukaryotic plankton; unless these novel eukaryotes are indeed a recent addition to the community. Biliphytes sequences were also found in most our clones libraries collected over a year in the Florida Straits. Seasonal clone libraries provide valuable information about phytoplankton populations as many studies focus on single time points. Phytoplankton communities are dynamic and the presence of biliphyte sequences, which are putatively photosynthetic, year-round at our study sites suggest that they are a recurrent component of the community, at least in the Florida Straits.

There are indications that biliphytes are photosynthetic, based on the phycobilin-like orange fluorescence observed under epifluorescence microscopy first reported by Not, Valentin and colleagues (2007). We observed such fluorescence as well. However, we did not observe any chlorophyll fluorescence in labeled cells. This is not surprising given that FISH samples are treated with an ethanol dehydration step and chlorophylls are pigments soluble in alcohol (Wright et al., 1997). In contrast phycobiliproteins that fluoresce orange under blue light epifluorescence microscopy are not eliminated during

the ethanol dehydration (Glazer, 1995, Not, Valentin et al., 2007). Nevertheless, the presence of a photosynthetic apparatus and ability to perform photosynthesis has still not been confirmed through classical characterization studies. An additional problem is the fact that only low numbers of small cells have been observed so far. Furthermore, the hypothesis that the orange fluorescence observed inside the cell could be a prey item such as *Synechococcus* cannot be excluded.

Biliphytes were also found to have a more extended ecological range than previously thought. Originally, this group was reported in cold waters from the Arctic Ocean, the Norwegian Sea and cold European coastal waters as well as from the permanent thermocline (500 m) near BATS (Romari and Vaulot, 2004, Not et al., 2007, Not, Valentin et al., 2007). We were able to detect cells in the Florida Straits in temperatures reaching 30 °C and retrieved sequences from clone libraries constructed from waters as cold as 5 °C on the continental shelf close to Woods Hole, MA. This indicates that biliphytes can survive in subtropical and tropical environments in addition to colder waters (see Not, Valentin et al., 2007), but no clear relationship was observed between water temperature and the phylogenetic clades. Future research on these organisms is needed to comprehensively address their importance. The distribution of biliphytes from the tropics to polar regions suggest that these organisms might be present in many ocean basins, but evidence will be needed to confirm this hypothesis. Furthermore, if biliphytes are indeed widely distributed, future studies will need to focus on their abundance and contribution to the carbon cycle. Are biliphytes key players in marine ecosystems or are they part scarce but highly diverse populations that constitute the rare biosphere (Sogin et al., 2006, Dawson and Hagen 2009). Additional questions remain about this group.

Evidence suggests the group specific probed cells were photosynthetic or based on their fluorescence, potentially mixotrophic (i.e. able to acquire energy and nutrients by phototrophic autotrophy as well as phagotrophic heterotrophy). In order to start understanding biliphyte roles in the food chain, further research is needed to address their trophic mode.

Uncultured prymnesiophytes were the focus of the third chapter of the dissertation. Prymnesiophytes are already known to be an important component of the phytoplankton in open oceans. HPLC often indicates that prymnesiophytes are in fact ubiquitous and dominate the eukaryotic phytoplankton (e.g. Steinberg et al., 2001). However, to date, almost all the prymnesiophytes species represented in culture collections belong to the nano-size fraction (3-20 μm, e.g. Vaulot et al., 2008, Worden and Not, 2008). Here we showed that a large portion of the environmental prymnesiophytes (i.e. with no cultured representative) border the pico-size fraction, nano-fraction split. Group-specific FISH revealed that *in situ* prymnesiophytes are small with average cell size of 3.4 ± 0.5 x 2.8 ± 0.6 μm (length x width). Using clone libraries constructed from <2-3 μm size-fractioned environmental samples this group is indeed extremely diverse. None of the sequences recovered in our libraries were strictly identical to known cultured prymnesiophytes and many sequences formed new phylogenetic clades. In addition, these small prymnesiophytes can contribute significantly to the total number of picoeukaryotic cells and form a significant portion of the picoplankton biomass in five major oceans. These findings expand upon earlier reports that environmentally relevant pico-prymnesiophytes are not represented in culture collections (Rappé et al., 1995, Moon van der Stay et al., 2000, Fuller et al., 2006, Worden and Not, 2008). From our experiments in the Sargasso

Sea, we can conclude that uncultured pico-prymnesiophytes have the potential for high growth rates. Therefore their contribution to primary production could be considerable, especially in the higher latitudes were they constitute a large portion of the picoplankton biomass. Despite this potential global importance, there is no knowledge yet of the distribution and abundance of specific clades or on ecological niches of these various clades. Some of the environmental pico-prymnesiophytes highlighted in this study (for example group 8) only share 93% identity at the nucleotide level (using 18S rRNA gene) to any cultured prymnesiophytes. This is less than, for example, two stains of *Micromonas* which have 97 % (18S rDNA) identity but only share 90% of their protein encoding genes (Worden et al., 2009). If such diversity is observed among the pico-prymnesiophytes, it must reflect differences in physiology and adaptations to specific environments. Future studies should address some of these potential adaptations and ecological niche differentiations. In addition, most of the pico-prymnesiophytes sequences recovered to date from environmental samples fall within the broad clade B2 which includes the genus *Chrysochromulina* and *Prymnesium* (order Prymnesiales). Many members of this order are mixotrophs (Nygaard and Tobiesen, 1993, Jones, 1994, Tillmann, 1998) suggesting that the success of pico-prymnesiophytes might be related to mixotrophy. The placement of many environmental sequences in the order Prymnesiales suggest that mixotrophy which is a common strategy at least for some of the larger prymnesiophytes could also be used by smaller uncultured prymnesiophytes. Heterotrophic flagellates are key players in controlling (cyano) bacterial (and potentially small eukaryotic populations), however, there is increasing evidence that small mixotrophic flagellates (<5 μm) are responsible for a significant portion of this

bacterivory as shown in the Mediterranean Sea and the Northern Atlantic (Unrein et al., 2007, Zubkov and Tarran, 2008). Although, the taxonomic affiliation of the mixotrophic flagellates was beyond the scope of the research, the authors suggest (based on microscopic observations in one study) that a substantial proportion of these small mixotrophs were prymnesiophytes (Unrein et al., 2007). Finally, evidence that prymnesiophytes are able to ingest *Prochlorococcus* was revealed by Frias-Lopez et al. (2009). This ability of pico-prymnesiophytes to change trophic modes depending on conditions can have significant implication in our understanding of the role of pico-prymnesiophytes in food webs and the carbon cycle. In the future, studies should address the extent of mixotrophy for this group, the potential impact on prey populations as well as the triggers that may determine the switch in trophic mode.

The fourth chapter provides a first look at light-related controls of a third important group of picoeukaryotes, the prasinophytes. In order to understand the dynamics observed in the environment, it is important to understand the mechanistic basis of organism response. Here, the response of *Micromonas* under high light (HL) and HL + UV radiation was studied. Flow cytometry analysis showed that cells exposed to both high light and UV radiation readjusted their chlorophyll content within the 2.5 h of treatment. The UV dosage ultimately was detrimental to the cultures. An additional goal was to look at the molecular responses that might relate to photoprotection. This was explored in a preliminary fashion by developing and applying gene probes in experiments similar to those exploring growth responses to light changes. The gene probes were designed to allow us to address hypotheses regarding several genes involved in light harvesting and thought to be involved in photoprotection.

Overall, this dissertation enhances our understanding of the distribution and abundance of uncultured picoeukaryotes. In general, knowledge focusing on specific picoplankton eukaryotic groups and their contributions is important for determining key players on local and global scales. While the relative importance of novel uncultured microbes to global primary production remains to be determined, their discovery allows reconsideration of some aspects of phytoplankton population dynamics. The wide diversity observed in both the pico-prymnesiophytes and biliphytes, as well as many other plankton groups, presumably must be linked to more fine scale niche partitioning and different functional roles. In the future, a combination of methods such as quantitative PCR (qPCR), FISH and other molecular approaches to target specific clades in their environments should be developed in order to address these questions. In addition, more studies should be geared toward understanding the factors that affect phytoplankton growth and physiology as well as how phylogenetic diversity translates into functional diversity. Exploration of microbial functional diversity is now possible through whole-genome sequencing of representatives of major taxa and genomics analysis of environmental samples. Approaches such as transcriptomics (the study of all the RNA transcripts present in a cell, population, or sample) and proteomics (the study of a complete set of a proteins from a cell, a population or an sample) on environmental samples, and on cultures, should help to expose novel proteins families and metabolic pathways and help shape hypothesis about key cellular activities or regulatory functions related to specific environments or perturbations. In the last few years the dramatic decrease in sequencing costs and development of new technologies such as next generation sequencing (e.g. SOLiD and Solexa) have offered new tools for such studies.

While environmental studies are valuable, laboratory experiments remain indispensable to test hypothesis in controlled environments. In addition, data such as that resulting from laboratory experiments (e.g. Chapter 4) can then prompt refined experiments in the field.

Overall, the diversity of marine microbial communities makes it difficult to determine which populations are most critical for input in global biogeochemical models. This work highlights the importance of understanding specific groups in the context of their environments in addition to experiments in the laboratory setting. Multidisciplinary approaches are critical to unveiling how organisms affect or are affected by their environment. Cross-disciplinary research should be emphasized, as it will allow a better understanding of the role of microbial population in marine ecosystems.

Literature Cited

Adamska, I. (1997). ELIPs: light-induced stress proteins. *Physiologia Plantarum*, *100*, 794-805.

Adamska, I., Kruse, E., & Kloppstech, K. (2001). Stable insertion of the early light-induced proteins into etioplast membranes requires chlorophyll *a*. *Journal of Biological Chemistry*, *276*(11), 8582-8587.

Adamska, I., Ohad, I., & Kloppstech, K. (1992). Synthesis of the early light-inducible protein is controlled by blue light and related to light stress. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(7), 2610-2613.

Adl, S. M., Simpson, A. G. B., Farmer, M. A., Andersen, R. A., Anderson, O. R., Barta, J. R., Bowser, S. S., et al. (2005). The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *The Journal of Eukaryotic Microbiology*, *52*(5), 399-451.

Alexander, E., Stock, A., Breiner, H., Behnke, A., Bunge, J., Yakimov, M. M., & Stoeck, T. (2009). Microbial eukaryotes in the hypersaline anoxic L'Atalante deep-sea basin. *Environmental Microbiology*, *11*(2), 360-381.

Andersen, R. A. (2004). Biology and systematics of heterokont and haptophyte algae. *American Journal of Botany*, *91*(10), 1508-1522.

Andersen, R. A., Bidigare, R. R., Keller, M. D., & Latasa, M. (1996). A comparison of HPLC pigment signatures and electron microscopic observations for oligotrophic waters of the North Atlantic and Pacific Oceans. *Deep Sea Research II*, *43*(2-3), 517-537.

Andersson, U., Heddad, M., & Adamska, I. (2003). Light stress-induced one-helix protein of the chlorophyll *a/b*-binding family associated with photosystem I. *Plant Physiology*, *132*(2), 811-820.

Anning, T., MacIntyre, H. L., Pratt, S. M., Sammes, P. J., Gibb, S., & Geider, R. J. (2000). Photoacclimation in the marine diatom *Skeletonema costatum*. *Limnology and Oceanography*, *45*(8), 1807-1817.

Archibald, J. M. (2007). Nucleomorph genomes: structure, function, origin and evolution. *BioEssays*, *29*(4), 392-402.

Azam, F., & Worden, A. Z. (2004). Microbes, molecules, and marine ecosystems. *Science*, *303*(5664), 1622-1624.

Baroli, I., Gutman, B. L., Ledford, H. K., Shin, J. W., Chin, B. L., Havaux, M., & Niyogi, K. K. (2004). Photo-oxidative stress in a xanthophyll-deficient mutant of *Chlamydomonas*. *Journal of Biological Chemistry*, *279*(8), 6337-6344.

Bei-Paraskevopoulou, T. & Kloppstech, K. (1999). The expression of early light-inducible proteins (ELIPs) under high-light stress as defense marker in Northern and Southern European cultivars of barley (*Hordeum vulgare*). *Physiologia Plantarum*, *106*(1), 105-111.

Bertilsson, S., Berglund, O., Karl, D. M., & Chisholm, S. W. (2003). Elemental composition of marine *Prochlorococcus* and *Synechococcus*: implications for the ecological stoichiometry of the sea. *Limnology and Oceanography*, *48*(5), 1721-1731.

Bidigare, R. R., & Ondrusek, M. E. (1996). Spatial and temporal variability of phytoplankton pigment distributions in the central equatorial Pacific Ocean. *Deep Sea Research II*, *43*(4-6), 809-833.

Binder, B. J., Chisholm, S. W., Olson, R. J., Frankel, S. L., & Worden, A. Z. (1996). Dynamics of picophytoplankton, ultraphytoplankton and bacteria in the central equatorial Pacific. *Deep Sea Research II*, *43*(4-6), 907-931.

Blanchot, J., André, J. -M., Navarette, C., Neveux, J., & Radenac, M. -H. (2001). Picophytoplankton in the equatorial Pacific: vertical distributions in the warm pool and in the high nutrient low chlorophyll conditions. *Deep Sea Research I*, *48*(1), 297-314.

Blanchot, J., & Rodier, M. (1996). Picophytoplankton abundance and biomass in the western tropical Pacific Ocean during the 1992 El Niño year: results from flow cytometry. *Deep Sea Research I*, *43*(6), 877-895.

Booth, B. C., Lewin, J., & Lorenzen, C. J. (1988). Spring and summer growth rates of subarctic Pacific phytoplankton assemblages determined from carbon uptake and cell volumes estimated using epifluorescence microscopy. *Marine Biology*, *98*(2), 287-298.

Booth, B. C., & Marchant, H. J. (1987). Parmales, a new order of marine chrysophytes with descriptions of three new genera and seven new species. *Journal of Phycology*, *23*(s2), 245-260.

Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., et al. (2008). The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*, *456*(7219), 239-244.

Bowler, C., Vardi, A., & Allen, A. E. (2010). Oceanographic and biogeochemical insights from diatom genomes. *Annual Review of Marine Science*, *2*(1), 333-365.

Brown, S. L., Landry, M. R., Barber, R. T., Campbell, L., Garrison, D. L., & Gowing, M. M. (1999). Picophytoplankton dynamics and production in the Arabian Sea during the 1995 Southwest Monsoon. *Deep Sea Research II*, *46*(8-9), 1745-1768.

Bruce, V. G. (1970). The biological clock in *Chlamydomonas reinhardtii*. *Journal of Protozoology*, *17*, 328-334.

Brunet, C., Casotti, R., Vantrepotte, V., & Conversano, F. (2007). Vertical variability and diel dynamics of picophytoplankton in the Strait of Sicily, Mediterranean Sea, in summer. *Marine Ecology Progress Series*, *346*, 15-26.

Buck, K. R., Chavez, F. P., & Campbell, L. (1996). Basin-wide distributions of living carbon components and the inverted trophic pyramid of the central gyre of the North Atlantic Ocean, summer 1993. *Aquatic Microbial Ecology*, *10*(3), 283-298.

Burki, F., Shalchian-Tabrizi, K., Minge, M., Skjæveland, Å., Nikolaev, S. I., Jakobsen, K. S., & Pawlowski, J. (2007). Phylogenomics reshuffles the eukaryotic supergroups. *PLoS ONE*, *2*(8), e790.

Cailliau, C., Claustre, H., Vidussi, F., Marie, D., & Vaulot, D. (1996). Carbon biomass, and gross growth rates as estimated from ^{14}C pigment labelling, during photoacclimation in *Prochlorococcus* CCMP 1378. *Marine Ecology Progress Series*, *145*, 209-221.

Campbell, L., Landry, M. R., Constantinou, J., Nolla, H. A., Brown, S. L., Liu, H., & Caron, D. A. (1998). Response of microbial community structure to environmental forcing in the Arabian Sea. *Deep Sea Research II*, *45*(10-11), 2301-2325.

Campbell, L., Nolla, H. A., & Vaulot, D. (1994). The Importance of *Prochlorococcus* to community structure in the central North Pacific Ocean. *Limnology and Oceanography*, *39*(4), 954-961.

Campbell, L., & Vaulot, D. (1993). Photosynthetic picoplankton community structure in the subtropical North Pacific Ocean near Hawaii (station ALOHA). *Deep Sea Research I*, *40*(10), 2043-2060.

Cardol, P., Bailleul, B., Rappaport, F., Derelle, E., Beal, D., Breyton, C., Bailey, S., et al. (2008). An original adaptation of photosynthesis in the marine green alga *Ostreococcus*. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(22), 7881-7886.

Carreto, J. I., Seguel, M., Montoya, N. G., Clement, A., & Carignan, M. O. (2001). Pigment profile of the ichthyotoxic dinoflagellate *Gymnodinium* sp. from a massive bloom in southern Chile. *Journal of Plankton Research*, *23*(10), 1171-1175.

Castenholz, R. W., & Garcia-Pichel, F. (2000). Cyanobacterial responses to UV-radiation. In: Whitton, B. A & Potts, M., Editors. Ecology of cyanobacteria: their diversity in time and space. Kluwer Academic Publishers, Dordrecht, the Netherlands. pp. 591-611.

Chapman, A. D. (2009). Numbers of living species in Australia and the world. Report for the Australian Biological Resources Study. Canberra, Australia. pp 1-82.

Chavez, F. P. (1989). Size distribution of phytoplankton in the central and eastern tropical Pacific. *Global Biogeochemical Cycles*, *3*(1), 27-35.

Cheung, M. K., Chu, K. H., Li, C. P., Kwan, H. S., & Wong, C. K. (2008). Genetic diversity of picoeukaryotes in a semi-enclosed harbour in the subtropical western Pacific Ocean. *Aquatic Microbial Ecology*, *53*(3), 295-305.

Chisholm, S. W. (1992). Phytoplankton size. In: Falkowski, P.G., & Woodhead, A.D. Editors. Primary production and biogeochemical cycles in the sea. Plenum Press, New York, New York, USA. pp: 213-237.

Chisholm, S. W., Frankel, S. L., Goericke, R., Olson, R. J., Palenik, B., Waterbury, J. B., West-Johnsrud, L., et al. (1992). *Prochlorococcus marinus* nov. gen. nov. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. *Archives of Microbiology*, *157*(3), 297-300.

Chisholm, S. W., Olson, R. J., Zettler, E. R., Goericke, R., Waterbury, J. B., & Welschmeyer, N. A. (1988). A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature*, *334*(6180), 340-343.

Chrétiennot-Dinet, M. -J., Sournia, A., Ricard, M., & Billard, C. (1993). A classification of the marine phytoplankton of the world from class to genus. *Phycologia*, *32*, 159-179.

Countway, P. D., Gast, R. J., Dennett, M. R., Savai, P., Rose, J. M., & Caron, D. A. (2007). Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). *Environmental Microbiology*, *9*(5), 1219-1232.

Countway, P. D., Gast, R. J., Savai, P., & Caron, D. A. (2005). Protistan diversity estimates based on 18S rDNA from seawater incubations in the Western North Atlantic. *The Journal of Eukaryotic Microbiology*, *52*(2), 95-106.

Cuvelier, M. L., Ortiz, A., Kim, E., Moehlig, H., Richardson, D. E., Heidelberg, J. F., Archibald, J. M., et al. (2008). Widespread distribution of a unique marine protistan lineage. *Environmental Microbiology*, *10*(6), 1621-1634.

Dawson, S., & Hagen, K. (2009). Mapping the protistan 'rare biosphere'. *Journal of Biology*, *8*(12), 105.

Dean, F. B., Hosono, S., Fang, L., Wu, X., Faruqi, A. F., Bray-Ward, P., Sun, Z., et al. (2002). Comprehensive human genome amplification using multiple displacement amplification. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(8), 5261-5266.

DeLong, E. F., & Karl, D. M. (2005). Genomic perspectives in microbial oceanography. *Nature*, *437*(7057), 336-342.

DeLong, E. F., Preston, C. M., Mincer, T., Rich, V., Hallam, S. J., Frigaard, N., Martinez, A., et al. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science*, *311*(5760), 496-503.

Derelle, E., Ferraz, C., Rombauts, S., Rouzé, P., Worden, A. Z., Robbens, S., Partensky, F., et al. (2006). Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(31), 11647-11652.

Díez, B., Pedós-Alió, C., & Massana, R. (2001). Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Applied and Environmental Microbiology*, *67*(7), 2932-2941.

Dolganov, N. A., Bhaya, D., & Grossman, A. R. (1995). Cyanobacterial protein with similarity to the chlorophyll *a/b* binding proteins of higher plants: evolution and regulation. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(2), 636-640.

Douglas, S., Zauner, S., Fraunholz, M., Beaton, M., Penny, S., Deng, L., Wu, X., et al. (2001). The highly reduced genome of an enslaved algal nucleus. *Nature*, *410*(6832), 1091-1096.

DuRand, M. D., & Olson, R. J. (1998). Diel patterns in optical properties of the chlorophyte *Nannochloris* sp.: Relating individual-cell to bulk measurements. *Limnology and Oceanography*, *43*(6), 1107-1118.

DuRand, M. D., Olson, R. J., & Chisholm, S. W. (2001). Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep Sea Research II*, *48*(8-9), 1983-2003.

Durnford, D. G., Price, J. A., McKim, S. M., & Sarchfield, M. L. (2003). Light-harvesting complex gene expression is controlled by both transcriptional and post-transcriptional mechanisms during photoacclimation in *Chlamydomonas reinhardtii*. *Physiologia Plantarum*, *118*(2), 193-205.

Edvardsen, B., Eikrem, W., Green, J. C., Andersen, R. A., Moon-van der Staay, S. Y., & Medlin, L. K. (2000). Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. *Phycologia*, *39*(1), 19-35.

Edvardsen, B., & Medlin, L. K. (2007). Molecular systematics of Haptophyta. In: Brodie, J., & Lewis, J. Editors. Unravelling the algae — the past, present and future of algal molecular systematics. Taylor and Francis, Boca Raton, Florida, USA. pp.183-196.

Eppley, R. W. (1968). An incubation method for estimating the carbon content of phytoplankton in natural samples. *Limnology and Oceanography*, *13*(4), 574-582.

Farinas, B., Mary, C., de O Manes, C., Bhaud, Y., Peaucellier, G., & Moreau, H. (2006). Natural synchronisation for the study of cell division in the green unicellular alga *Ostreococcus tauri*. *Plant Molecular Biology*, *60*(2), 277-292.

Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package) version 3.6. *Distributed by the author.* Department of Genome Sciences, University of Washington, Seattle.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, *281*(5374), 237-240.

Fratantoni, P. S., Lee, T. N., Podesta, G. P., & Muller-Karger, F. (1998). The influence of Loop Current perturbations on the formation and evolution of Tortugas eddies in the southern Straits of Florida. *Journal of Geophysical Research*, *103*(C11), PP. 24,759-24,779.

Frias-Lopez, J., Thompson, A., Waldbauer, J., & Chisholm, S. W. (2009). Use of stable isotope-labelled cells to identify active grazers of picocyanobacteria in ocean surface waters. *Environmental Microbiology*, *11*(2), 512-525.

Fujiwara, S., Tsuzuki, M., Kawachi, M., Minaka, N., & Inouye, I. (2001). Molecular phylogeny of the Haptophyta based on the *rbc*L gene and sequence variation in the spacer region of the RUBISCO operon. *Journal of Phycology*, *37*(1), 121-129.

Fuller, N. J., Campbell, C., Allen, D. J., Pitt, F. D., Zwirglmaier, K., Gall, F. L., Vaulot, D., et al. (2006). Analysis of photosynthetic picoeukaryote diversity at open ocean sites in the Arabian Sea using a PCR biased towards marine algal plastids. *Aquatic Microbial Ecology*, *43*(1), 79-93.

Gagné, G., & Guertin, M. (1992). The early genetic response to light in the green unicellular alga *Chlamydomonas eugametos* grown under light/dark cycles involves genes that represent direct responses to light and photosynthesis. *Plant Molecular Biology*, *18*(3), 429-445.

Garczarek, L., Partensky, F., Irlbacher, H., Holtzendorff, J., Babin, M., Mary, I., Thomas, J. C., et al. (2001). Differential expression of antenna and core genes in *Prochlorococcus* PCC 9511 (Oxyphotobacteria) grown under a modulated light-dark cycle. *Environmental Microbiology*, *3*(3), 168-175.

Glazer, A. (1994). Phycobiliproteins — a family of valuable, widely used fluorophores. *Journal of Applied Phycology*, *6*(2), 105-112.

Goericke, R., & Welschmeyer, N. A. (1998). Response of Sargasso Sea phytoplankton biomass, growth rates and primary production to seasonally varying physical forcing. *Journal of Plankton Research*, *20*(12), 2223-2249.

Green, B. R., Anderson, J. M., & Parson, W. W. (2003). Photosynthetic membranes and their light-harvesting antennas. In: Green B. R. & Parson W. W., Editors. Advances in photosynthesis and respiration - Light-harvesting antennas in photosynthesis. Vol. 13. Kluwer Academic Publishers, Dordrecht, the Netherlands. pp. 1-28.

Green, J. C., & Jordan, R. W. (1994). Systematic history and taxonomy. In: Green, J.C., and Leadbeater, B.S.C., Editors. The haptophyte algae. Oxford University Press, Oxford, UK. pp. 1-21.

Green, B. R., & Kühlbrandt, W. (1995). Sequence conservation of light-harvesting and stress-response proteins in relation to the three-dimensional molecular structure of LHCII. *Photosynthesis Research*, *44*(1), 139-148.

Grob, C., Ulloa, O., Li, W. K. W., Alarcn, G., Fukasawa, M., & Watanabe, S. (2007). Picoplankton abundance and biomass across the eastern South Pacific Ocean along latitude 32.5°S. *Marine Ecology Progress Series*, *332*, 53-62.

Groisillier, A., Massana, R., Valentin, K., Vaulot, D., & Guillou, L. (2006). Genetic diversity and habitats of two enigmatic marine alveolate lineages. *Aquatic Microbial Ecology*, *42*(3), 277-291.

Guillou, L., Chrétiennot-Dinet, M.-J., Medlin, L. K., Claustre, H., Loiseaux-de Goer, S., & Vaulot, D. (1999). *Bolidomonas*: a new genus with two species belonging to a new algal class, the Bolidophyceae (Heterokonta). *Journal of Phycology*, *35*(2), 368-381.

Guillou, L., Eikrem, W., Chrétiennot-Dinet, M.-J., Le Gall, F., Massana, R., Romari, K., Pedrós-Alió, C., et al. (2004). Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. *Protist*, *155*(2), 193-214.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*(5), 696-704.

Hampl, V., Hug, L., Leigh, J. W., Dacks, J. B., Lang, B. F., Simpson, A. G. B., & Roger, A. J. (2009). Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". *Proceedings of the National Academy of Sciences of the United States of America*, *106*(10), 3859-3864.

Harari-Steinberg, O., Ohad, I., & Chamovitz, D. A. (2001). Dissection of the light signal transduction pathways regulating the two *early light-induced protein* genes in Arabidopsis. *Plant Physiology*, *127*(3), 986-997.

Harris, G. N., Scanlan, D. J., & Geider, R. J. (2009). Responses of *Emiliania huxleyi* (Prymnesiophyceae) to step changes in photon flux density. *European Journal of Phycology*, *44*(1), 31.

He, Q., Dolganov, N., Björkman, O., & Grossman, A. R. (2001). The high light-inducible polypeptides in *Synechocystis* PCC6803. *Journal of Biological Chemistry*, *276*(1), 306-314.

Hearn, K. (2007). Bizarre new form of life found in Arctic Ocean, scientists announce. *National Geographic News*: National Geographic.

Heddad, M., & Adamska, I. (2000). Light stress-regulated two-helix proteins in *Arabidopsis thaliana* related to the chlorophyll *a/b*-binding gene family. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(7), 3741-3746.

Heddad, M., & Adamska, I. (2002). The evolution of light stress proteins in photosynthetic organisms. *Comparative and Functional Genomics*, *3*(6), 504-510.

Heddad, M., Noren, H., Reiser, V., Dunaeva, M., Andersson, B., & Adamska, I. (2006). Differential expression and localization of early light-induced proteins in Arabidopsis. *Plant Physiology.*, *142*(1), 75-87.

Hsiao, H., He, Q., van Waasbergen, L. G., & Grossman, A. R. (2004). Control of photosynthetic and high-light-responsive genes by the histidine kinase DspA: negative and positive regulation and interactions between signal transduction pathways. *Journal of Bacteriology*, *186*(12), 3882-3888.

Huang, L., McCluskey, M. P., Ni, H., & LaRossa, R. A. (2002). Global gene expression profiles of the cyanobacterium *Synechocystis* sp. strain PCC 6803 in response to irradiation with UV-B and white light. *Journal of Bacteriology*, *184*(24), 6845-6858.

Hutin, C., Nussaume, L., Moise, N., Moya, I., Kloppstech, K., & Havaux, M. (2003). Early light-induced proteins protect *Arabidopsis* from photooxidative stress. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(8), 4921-4926.

Jacquet, S., Partensky, F., Lennon, J., & Vaulot, D. (2001). Diel patterns of growth and division in marine picoplankton in culture. *Journal of Phycology*, *37*(3), 357-369.

Jansson, S. (1999). A guide to the *Lhc* genes and their relatives in *Arabidopsis*. *Trends in Plant Science*, *4*(6), 236-240.

Jansson, S., Andersson, J., Jung Kim, S., & Jackowski, G. (2000). An *Arabidopsis thaliana* protein homologous to cyanobacterial high-light-inducible proteins. *Plant Molecular Biology*, *42*(2), 345-351.

Johnson, P. W., & Sieburth, J. M. (1982). In-situ morphology and occurrence of eukaryotic phototrophs of bacterial size in the picoplankton of estuarine and oceanic waters. *Journal of Phycology*, *18*(3), 318-327.

Johnson, Z. I., Zinser, E. R., Coe, A., McNulty, N. P., Woodward, E. M. S., & Chisholm, S. W. (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science*, *311*(5768), 1737-1740.

Jones, H. L. J. (1994). Mixotrophy in haptophytes. In: Green, J. C., & Leadbeater, B. S.C. Editors. The haptophyte algae, The Systematics Association Special Volume No 51. Clarendon Press, Oxford, UK. pp 247-263.

Jordan, R. W., Cros, L., & Young, J. R. (2004). A revised classification scheme for living haptophytes. *Micropaleontology*, *50*, 55-79.

Kana, T. M., & Glibert, P. M. (1987). Effect of irradiances up to 2000 μE m^{-2} s^{-1} on marine *Synechococcus* WH7803. Growth, pigmentation, and cell composition. *Deep Sea Research A*, *34*(4), 479-495.

Keeling, P. J., Burger, G., Durnford, D. G., Lang, B. F., Lee, R. W., Pearlman, R. E., Roger, A. J., A. J. Roger & M. W. Gray (2005). The tree of eukaryotes. *Trends in Ecology & Evolution*, *20*(12), 670-676.

Keller, M. D., Selvin, R. C., Claus, W., & Guillard, R. R. L. (1987). Media for the culture of oceanic ultraphytoplankton. *Journal of Phycology*, *23*(4), 633-638.

Kimura, M., Yamamoto, Y. Y., Seki, M., Sakurai, T., Sato, M., Abe, T., Yoshida, S., et al. (2003). Identification of *Arabidopsis* genes regulated by high light-stress using cDNA microarray. *Photochemistry and Photobiology*, *77*(2), 226-233.

Kirk, J. T. O. (1994). Optics of UV-B radiation in natural waters. *Archiv für Hydrobiologie Beiheft Ergebnisse der Limnologie*, *43*, 1-16.

Klein Breteler, W. C. M. (1985). Fixation artifacts of phytoplankton in zooplankton grazing experiments. *Aquatic Ecology*, *19*(1), 13-19.

Kloppstech, K. (1997). Light regulation of photosynthetic genes. *Physiologia Plantarum*, *100*(4), 739-747.

Koziol, A. G., Borza, T., Ishida, K., Keeling, P., Lee, R. W., & Durnford, D. G. (2007). Tracing the evolution of the light-harvesting antennae in chlorophyll *a/b*-containing organisms. *Plant Physiology*, *143*(4), 1802-1816.

Kucho, K., Okamoto, K., Tabata, S., Fukuzawa, H., & Ishiura, M. (2005). Identification of novel clock-controlled genes by cDNA macroarray analysis in *Chlamydomonas reinhardtii*. *Plant Molecular Biology*, *57*(6), 889-906.

Landry, M. R. (1993). Estimating rates of growth and grazing mortality of phytoplankton by the dilution method. In: Kemp, P. F., Sherr, B. F., Sherr, E. B., and Cole, J. J., Editors. In: handbook of methods in aquatic microbial ecology. Lewis Publishers, Boca Raton, Florida, USA. pp. 715-722.

Landry, M. R., Brown, S. L., Neveux, J., Dupouy, C., Blanchot, J., Christensen, S., & Bidigare, R. R. (2003). Phytoplankton growth and microzooplankton grazing in high-nutrient, low-chlorophyll waters of the equatorial Pacific: community and taxon-specific rate assessments from pigment and flow cytometric analyses. *Journal of Geophysical Research*, *108,* 8142, doi:10.1029/2000JC000744.

Landry, M. R., & Hassett, R. P. (1982). Estimating the grazing impact of marine micro-zooplankton. *Marine Biology*, *67*(3), 283-288.

Landry, R, M., Kirshtein, J, Constantinou, & J. (1995). A refined dilution technique for measuring the community grazing impact of microzooplankton, with experimental tests in the central equatorial Pacific. *Marine Ecology Progress Series*, *120*, 53-63.

Landry, M. R., Kirshtein, J., & Constantinou, J. (1996). Abundances and distributions of picoplankton populations in the central equatorial Pacific from 12°N to 12°S, 140°W. *Deep Sea Research II*, *43*(4-6), 871-890.

Lange, M., Chen, Y. Q., & Medlin, L. K. (2002). Molecular genetic delineation of *Phaeocystis* species (Prymnesiophyceae) using coding and non-coding regions of nuclear and plastid genomes. *European Journal of Phycology*, *37*(01), 77-92.

Latasa, M. (2007). Improving estimations of phytoplankton class abundances using CHEMTAX. *Marine Ecology Progress Series*, *329*, 13-21.

Latasa, M., Scharek, R., Gall, F. L., & Guillou, L. (2004). Pigment suites and taxonomic groups in Prasinophyceae. *Journal of Phycology*, *40*(6), 1149-1155.

Latasa, M., van Lenning, K., Garrido, J., Scharek, R., Estrada, M., Rodríguez, F., & Zapata, M. (2001). Losses of chlorophylls and carotenoids in aqueous acetone and methanol extracts prepared for RPHPLC analysis of pigments. *Chromatographia*, *53*(7), 385-391.

Ledford, H. K., Baroli, I., Shin, J. W., Fischer, B. B., Eggen, R. I. L., & Niyogi, K. K. (2004). Comparative profiling of lipid-soluble antioxidants and transcripts reveals two phases of photo-oxidative stress in a xanthophyll-deficient mutant of *Chlamydomonas reinhardtii*. *Molecular Genetics and Genomics*, *272*(4), 470-479.

Lee, T. N., Clarke, M., Williams, E., Szmant, A. F., & Berger, T. (1994). Evolution of the Tortugas gyre and its influence on recruitment in the Florida Keys. *Bulletin of Marine Science*, *54*, 621-646.

Letelier, R. M., Bidigare, R. R., Hebel, D. V., Ondrusek, M., Winn, C. D., & Karl, D. M. (1993). Temporal variability of phytoplankton community structure based on pigment analysis. *Limnology and Oceanography*, *38*(7), 1420-1437.

Li, W. K. W. (1994). Primary production of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: measurements from flow cytometric sorting. *Limnology and Oceanography*, *39*(1), 169-175.

Li, W. K. W., Dickie, P., Irwin, B., & Wood, A. (1992). Biomass of bacteria, cyanobacteria, prochlorophytes and photosynthetic eukaryotes in the Sargasso Sea. *Deep Sea Research A, 39*(3-4), 501-519.

Li, W. K., Harrison, W. G., & Head, E. J. (2006). Coherent assembly of phytoplankton communities in diverse temperate ocean ecosystems. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1596), 1953-1960.

Li, W. K. W., McLaughlin, F. A., Lovejoy, C., & Carmack, E. C. (2009). Smallest algae thrive as the Arctic Ocean freshens. *Science*, *326*(5952), 539.

Liu, H., Probert, I., Uitz, J., Claustre, H., Aris-Brosou, S., Frada, M., Not, F., et al. (2009). Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(31), 12803-12808.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}{}_T$ method. *Methods*, *25*(4), 402-408.

López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C., & Moreira, D. (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, *409*(6820), 603-607.

Lovejoy, C., Massana, R., & Pedrós-Alió, C. (2006). Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent seas. *Applied and Environmental Microbiology*, *72*(5), 3085-3095.

Lovejoy, C., Vincent, W. F., Bonilla, S., Roy, S., Martineau, M., Terrado, R., Potvin, M., et al. (2007). Distribution, phylogeny, and cold-adapted picoprasinophytes in Arctic seas. *Journal of Phycology*, *43*(1), 78-89.

Ludwig, M., & Gibbs, S. P. (1987). Are the nucleomorphs of cryptomonads and *Chlorarachnion* the vestigial nuclei of eukaryotic endosymbionts? *Annals of the New York Academy of Sciences*, *503*, 198-211.

MacIntyre, H. L., Kana, T. M., & Geider, R. J. (2000). The effect of water motion on short-term rates of photosynthesis by marine phytoplankton. *Trends in Plant Science*, *5*(1), 12-17.

Mackey, M., Mackey, D., Higgins, H., & Wright, S. (1996). CHEMTAX - a program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton. *Marine Ecology Progress Series*, *144*, 265-283.

Mackey, D. J., Blanchot, J., Higgins, H. W., & Neveux, J. (2002). Phytoplankton abundances and community structure in the equatorial Pacific. *Deep Sea Research II*, *49*(13-14), 2561-2582.

Maidak, B. L., Cole, J. R., Lilburn, T. G., Parker, C. T., Saxman, P. R., Farris, R. J., Garrity, G. M., et al. (2001). The RDP-II (Ribosomal Database Project). *Nucleic Acids Research*, *29*(1), 173-174.

Marañón, E., Behrenfeld, M. J., González, N., Mouriño, B., & Zubkov, M. V. (2003). High variability of primary production in oligotrophic waters of the Atlantic Ocean: uncoupling from phytoplankton biomass and size structure. *Marine Ecology Progress Series*, *257*, 1-11.

Marie, D., Zhu, F., Balagué, V., Ras, J., & Vaulot, D. (2006). Eukaryotic picoplankton communities of the Mediterranean Sea in summer assessed by molecular approaches (DGGE, TTGE, QPCR). *FEMS Microbiology Ecology*, *55*(3), 403-415.

Massana, R., Balagué, V., Guillou, L., & Pedrós-Alió, C. (2004a). Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. *FEMS Microbiology Ecology*, *50*(3), 231-243.

Massana, R., Castresana, J., Balagué, V., Guillou, L., Romari, K., Groisillier, A., Valentin, K., et al. (2004b). Phylogenetic and ecological analysis of novel marine stramenopiles. *Applied and Environmental Microbiology*, *70*(6), 3528-3534.

Massana, R., & Pedrós-Alió, C. (2008). Unveiling new microbial eukaryotes in the surface ocean. *Current Opinion in Microbiology*, *11*(3), 213-218.Moon-van der Staay, S. Y., van der Staay, G. W. M., Guillou, L., Vaulot, D., Claustre, H., & Medlin, L. K. (2000). Abundance and diversity of prymnesiophytes in the picoplankton community from the equatorial Pacific Ocean inferred from 18S rDNA sequences. *Limnology and Oceanography*, *45*(1), 98-109.

Masuda, T., Tanaka, A., & Melis, A. (2003). Chlorophyll antenna size adjustments by irradiance in *Dunaliella salina* involve coordinate regulation of chlorophyll *a* oxygenase (*CAO*) and *Lhcb* gene expression. *Plant Molecular Biology*, *51*(5), 757-771.

McDonald, S. M., Plant, J. N., & Worden, A. Z. (2010). The mixed lineage nature of nitrogen transport and assimilation in marine eukaryotic phytoplankton: a case study of *Micromonas*. *Molecular Biology Evolution*, msq113.

McGillicuddy, D. J., Anderson, L. A., Bates, N. R., Bibby, T., Buesseler, K. O., Carlson, C. A., Davis, C. S., et al. (2007). Eddy/wind interactions stimulate extraordinary mid-ocean plankton blooms. *Science*, *316*(5827), 1021-1026.

Meador, J., Jeffrey, W. H., Kase, J. P., Pakulski, J. D., Chiarello, S., & Mitchell, D. L. (2002). Seasonal fluctuation of DNA photodamage in marine plankton assemblages at Palmer Station, Antarctica. *Photochemistry and Photobiology*, *75*(3), 266-271.

Medlin, L., Sáez, A., & Young, J. (2008). A molecular clock for coccolithophores and implications for selectivity of phytoplankton extinctions across the K/T boundary. *Marine Micropaleontology*, *67*(1-2), 69-86.

Medlin, L., & Zingone, A. (2007). A taxonomic review of the genus *Phaeocystis*. *Biogeochemistry*, *83*(1), 3-18.

Meyer, G., & Kloppstech, K. (1984). A rapidly light-induced chloroplast protein with a high turnover coded for by pea nuclear DNA. *European Journal of Biochemistry*, *138*(1), 201-207.

Miura, K., Yamano, T., Yoshioka, S., Kohinata, T., Inoue, Y., Taniguchi, F., Asamizu, E., et al. (2004). Expression profiling-based identification of CO_2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Physiology*, *135*(3), 1595-1607.

Monger, B. C., & Landry, M. R. (1993). Flow cytometric analysis of marine bacteria with Hoechst 33342. *Applied and Environmental Microbiology*, *59*(3), 905-911.

Monnier, A., Liverani, S., Bouvet, R., Jesson, B., Smith, J., Mosser, J., Corellou, F., et al. (2010). Orchestrated transcription of biological processes in the marine picoeukaryote *Ostreococcus* exposed to light/dark cycles. *BMC Genomics*, *11*(1), 192.

Montané, M. H., & Kloppstech, K. (2000). The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function? *Gene*, *258*(1-2), 1-8.

Moon-van der Staay, S. Y., De Wachter, R., & Vaulot, D. (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature*, *409*(6820), 607-610.

Moon-van der Staay, S. Y., van der Staay, G. W. M., Guillou, L., Vaulot, D., Claustre, H., & Medlin, L. K. (2000). Abundance and diversity of prymnesiophytes in the picoplankton community from the equatorial Pacific Ocean inferred from 18S rDNA sequences. *Limnology and Oceanography*, *45*(1), 98-109.

Moore, L. R., Rocap, G., & Chisholm, S. W. (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature*, *393*(6684), 464-467.

Moran, M. A., & Miller, W. L. (2007). Resourceful heterotrophs make the most of light in the coastal ocean. *Nature Reviews Microbiology*, *5*(10), 792-800.

Moulager, M., Monnier, A., Jesson, B., Bouvet, R., Mosser, J., Schwartz, C., Garnier, L., et al. (2007). Light-dependent regulation of cell division in *Ostreococcus*: evidence for a major transcriptional input. *Plant Physiology*, *144*(3), 1360-1369.

Nahon, S., Charles, F., & Pruski, A. M. (2008). Improved Comet assay for the assessment of UV genotoxicity in Mediterranean Sea urchin eggs. *Environmental and Molecular Mutagenesis*, *49*(5), 351-359.

Naumann, B., Busch, A., Allmer, J., Ostendorf, E., Zeller, M., Kirchhoff, H., & Hippler, M. (2007). Comparative quantitative proteomics to investigate the remodeling of bioenergetic pathways under iron deficiency in *Chlamydomonas reinhardtii*. *Proteomics*, *7*(21), 3964-3979.

Niyogi, K. K. (1999). Photoprotection revised: genetic and molecular approaches. *Annual Review of Plant Physiology and Plant Molecular Biology*, *50*, 333-359.

Norén, H., Svensson, P., Stegmark, R., Funk, C., Adamska, I., & Andersson, B. (2003). Expression of the early light-induced protein but not the PsbS protein is influenced by low temperature and depends on the developmental stage of the plant in field-grown pea cultivars. *Plant, Cell & Environment*, *26*(2), 245-253.

Not, F., Gausling, R., Azam, F., Heidelberg, J. F., & Worden, A. Z. (2007). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environmental Microbiology*, *9*(5), 1233-1252.

Not, F., Latasa, M., Scharek, R., Viprey, M., Karleskind, P., Balagué, V., Ontoria-Oviedo, I., et al. (2008). Protistan assemblages across the Indian Ocean, with a specific emphasis on the picoeukaryotes. *Deep Sea Research I*, *55*(11), 1456-1473.

Not, F., Simon, N., Biegala, I. C., & Vaulot, D. (2002). Application of fluorescent *in situ* hybridization coupled with tyramide signal amplification (FISH-TSA) to assess eukaryotic picoplankton composition. *Aquatic Microbial Ecology*, *28*(2), 157-166.

Not, F., Valentin, K., Romari, K., Lovejoy, C., Massana, R., Töbe, K., Vaulot, D., & Medlin, L. K. (2007). Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. *Science*, *315*(5809), 253-255.

Nygaard, K., & Tobiesen, A. (1993). Bacterivory in algae: a survival strategy during nutrient limitation. *Limnology and Oceanography*, *38*(2), 273-279.

Olson, R. J., Chisholm, S. W., Zettler, E. R., Altabet, M. A., & Dusenberry, J. A. (1990). Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. *Deep Sea Research A*, *37*(6), 1033-1051.

Olson, R. L., Zettler, E. R., and DuRand, M. D. (1993). Phytoplankton analysis using flow cytometry. In: Kemp, P. F., Sherr, B. F., Sherr, E. B., & Cole, J. J. Editors. Handbook of methods in aquatic ecology. Lewis Publishers, Boca Raton, Florida, USA. pp. 175-186.

Palenik, B., Grimwood, J., Aerts, A., Rouzé, P., Salamov, A., Putnam, N., Dupont, C., et al. (2007). The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(18), 7705-7710.

Palmer, J. D. (2003). The Symbiotic birth and spread of plastids: how many times and whodunit? *Journal of Phycology*, *39*(1), 4-12.

Parke, M. & Dixon P. S. (1976). Check-list of British marine algae-third revision. *Journal of the Marine Biological Association of the United Kingdom 56*, 527-594.

Parker, M. S., Mock, T., & Armbrust, E. V. (2008). Genomic insights into marine microalgae. *Annual Review of Genetics*, *42*, 619-645.

Partensky, F., Blanchot, J., Lantoine, F., Neveux, J., & Marie, D. (1996). Vertical structure of picophytoplankton at different trophic sites of the tropical northeastern Atlantic Ocean. *Deep Sea Research I, 43*(8), 1191-1213.

Partensky, F., Hess, W. R., & Vaulot, D. (1999). *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiology and Molecular Biology Reviews*, *63*(1), 106-127.

Patron, N. J., Inagaki, Y., & Keeling, P. J. (2007). Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. *Current Biology*, *17*(10), 887-891.

Paul, J. H., Alfreider, A., & Wawrik, B. (2000). Micro- and macrodiversity in *rbcL* sequences in ambient phytoplankton populations from the southeastern Gulf of Mexico. *Marine Ecology Progress Series*, *198*, 9-18.

Peers, G., Truong, T. B., Ostendorf, E., Busch, A., Elrad, D., Grossman, A. R., Hippler, M., et al. (2009). An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature*, *462*(7272), 518-521.

Pérez, V., Fernández, E., Marañón, E., Serret, P., Varela, R., Bode, A., Varela, M., et al. (2005). Latitudinal distribution of microbial plankton abundance, production, and respiration in the Equatorial Atlantic in autumn 2000. *Deep Sea Research I*, *52*(5), 861-880.

Posada, D., & Crandall, K. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics*, *14*(9), 817-818.

Post, A. F., Dubinsky, Z., Wyman, K., & Falkowski, P. G. (1984). Kinetics of light-intensity adaptation in a marine planktonic diatom. *Marine Biology*, *83*(3), 231-238.

Rappé, M. S., Kemp, P. F., & Giovannoni, S. J. (1995). Chromophyte plastid 16S ribosomal RNA genes found in a clonal library from Atlantic Ocean seawater. *Journal of Phycology*, *31*(6), 979-988.

Rappé, M. S., Suzuki, M. T., Vergin, K. L., & Giovannoni, S. J. (1998). Phylogenetic diversity of ultraplankton plastid small-subunit rRNA genes recovered in environmental nucleic acid samples from the Pacific and Atlantic coasts of the United States. *Applied and Environmental Microbiology*, *64*(1), 294-303.

Raven, J. A. (1986). Physiological consequences of extremely small size for autotrophic organisms in the sea. *Canadian Bulletin of Fisheries and Aquatic Sciences*, *214*, 1-70.

Raven, J. A. (1998). The twelfth Tansley Lecture. Small is beautiful: the picophytoplankton. *Functional Ecology*, *12*(4), 503-513.

Richard, C., Ouellet, H., & Guertin, M. (2000). Characterization of the LI818 polypeptide from the green unicellular alga *Chlamydomonas reinhardtii*. *Plant Molecular Biology*, *42*(2), 303-316.

Richardson, T. L., & Jackson, G. A. (2007). Small phytoplankton and carbon export from the surface ocean. *Science*, *315*(5813), 838-840.

Rocap, G., Larimer, F. W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N. A., Arellano, A., et al. (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature*, *424*(6952), 1042-1047.

Rodríguez, F., Derelle, E., Guillou, L., Gall, F. L., Vaulot, D., & Moreau, H. (2005). Ecotype diversity in the marine picoeukaryote *Ostreococcus* (Chlorophyta, Prasinophyceae). *Environmental Microbiology*, *7*(6), 853-859.

Romari, K., & Vaulot, D. (2004). Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. *Limnology and Oceanography*, *19*(3), 784-798.

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., Wu, D., et al. (2007). The Sorcerer II Global Ocean Sampling Expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biology*, *5*(3), e77.

Sáez, A., Probert, I., Young, J., Edvardsen , B., Eikrem, W., and Medlin, L. K. (2004). A review of the phylogeny of the Haptophyta. In: Thierstein, H. & Young, J., Editors. Coccolithophores—from molecular processes to global impact. Springer, Berlin, Germnay. pp. 251-269.

Salem, K., & van Waasbergen, L. G. (2004). Light control of *hliA* transcription and transcript stability in the cyanobacterium *Synechococcus elongatus* strain PCC 7942. *Journal of Bacteriology*, *186*(6), 1729-1736.

Savard, F., Richard, C., & Guertin, M. (1996). The *Chlamydomonas reinhardtii LI818* gene represents a distant relative of the *cabI/II* genes that is regulated during the cell cycle and in response to illumination. *Plant Molecular Biology*, *32*(3), 461-473.

Shalapyonok, A., Olson, R. J., & Shalapyonok, L. S. (2001). Arabian Sea phytoplankton during Southwest and Northeast Monsoons 1995: composition, size structure and biomass from individual cell properties measured by flow cytometry. *Deep Sea Research II*, *48*(6-7), 1231-1261.

Shi, X. L., Marie, D., Jardillier, L., Scanlan, D. J., & Vaulot, D. (2009). Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic south east Pacific Ocean. *PLoS ONE*, *4*(10), e7657.

Sieburth, J. M., Smetacek, V., & Lenz, J. (1978). Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnology and Oceanography*, *23*(6), 1256-1263.

Simmons, M. P., Pickett, K. M., & Miya, M. (2004). How meaningful are Bayesian support values? *Molecular Biology Evolution*, *21*(1), 188-199.

Simon, N., Campbell, L., Örnólfsdóttir, E., Groben, R., Guillou, L., Lange, M., & Medlin, L. K. (2000). Oligonucleotide probes for the identification of three algal groups by dot blot and fluorescent whole-cell hybridization. *The Journal of Eukaryotic Microbiology*, *47*(1), 76-84.

Simon, N., Cras, A., Foulon, E., & Lemée, R. (2009). Diversity and evolution of marine phytoplankton. *Comptes Rendus Biologies*, *332*(2-3), 159-170.

Six, C., Worden, A. Z., Rodríguez, F., Moreau, H., & Partensky, F. (2005). New insights into the nature and phylogeny of prasinophyte antenna proteins: *Ostreococcus tauri*, a case study. *Molecular Biology and Evolution*, *22*(11), 2217-2230.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J. M., et al. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America*, *103*(32), 12115-12120.

Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., & Knap, A. H. (2001). Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. *Deep Sea Research II*, *48*(8-9), 1405-1447.

Stockner, J. G. (1988). Phototrophic picoplankton: an overview from marine and freshwater ecosystems. *Limnology and Oceanography*, *33*(4), 765-775.

Takano, Y., Hagino, K., Tanaka, Y., Horiguchi, T., & Okada, H. (2006). Phylogenetic affinities of an enigmatic nannoplankton, *Braarudosphaera bigelowii* based on the SSU rDNA sequences. *Marine Micropaleontology*, *60*(2), 145-156.

Tedetti, M., & Sempéré, R. (2006). Penetration of ultraviolet radiation in the marine environment. A review. *Photochemistry and Photobiology*, *82*(2), 389-397.

Teramoto, H., Itoh, T., & Ono, T. (2004). High-intensity-light-dependent and transient expression of new genes encoding distant relatives of light-harvesting chlorophyll-*a/b* proteins in *Chlamydomonas reinhardtii*. *Plant Cell Physiology*, *45*(9), 1221-1232.

Teramoto, H., Nakamori, A., Minagawa, J., & Ono, T. (2002). Light-intensity-dependent expression of *Lhc* gene family encoding light-harvesting chlorophyll-*a/b* proteins of photosystem II in *Chlamydomonas reinhardtii*. *Plant Physiology*, *130*(1), 325-333.

Tevini, M., Braun, J., & Fieser, G. (1991). The protective function of the epidermal layer of rye seedlings against ultraviolet-B radiation. *Photochemistry and Photobiology*, *53*(3), 329-333.

Tillmann, U. (1998). Phagotrophy by a plastidic haptophyte, *Prymnesium patelliferum*. *Aquatic Microbial Ecology*, *14*(2), 155-160.

Unrein, F., Massana, R., Alonso-Sáez, L., & Gasol, J. M. (2007). Significant year-round effect of small mixotrophic flagellates on bacterioplankton in an oligotrophic coastal system. *Limnology and Oceanography*, *52*(1), 456-469.

Van Mooy, B. A. S., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblizek, M., Lomas, M. W., et al. (2009). Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature*, *458*(7234), 69-72.

van Waasbergen, L. G., Dolganov, N., & Grossman, A. R. (2002). *nblS*, a gene involved in controlling photosynthesis-related gene expression during high light and nutrient stress in *Synechococcus elongatus* PCC 7942. *Journal of Bacteriology*, *184*(9), 2481-2490.

Vaulot, D., Eikrem, W., Viprey, M., & Moreau, H. (2008). The diversity of small eukaryotic phytoplankton (≤ 3 μm) in marine ecosystems. *FEMS Microbiology Reviews*, *32*(5), 795-820.

Viprey, M., Guillou, L., Ferréol, M., & Vaulot, D. (2008). Wide genetic diversity of picoplanktonic green algae (Chloroplastida) in the Mediterranean Sea uncovered by a phylum-biased PCR approach. *Environmental Microbiology*, *10*(7), 1804-1822.

Waterbury, J. B., Watson, S., Guillard, R. R. L., & Brand, L. E. (1979). Widespread occurrence of a unicellular, marine, planktonic, cyanobacterium. *Nature*, *277*(5694), 293-294.

Worden, A. Z. (2006). Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquatic Microbial Ecology*, *43*(2), 165-175.

Worden, A. Z., & Binder, B. J. (2003). Application of dilution experiments for measuring growth and mortality rates among *Prochlorococcus* and *Synechococcus* populations in oligotrophic environments. *Aquatic Microbial Ecology*, *30*(2), 159-174.

Worden, A. Z., Chisholm, S. W., & Binder, B. J. (2000). In situ hybridization of *Prochlorococcus* and *Synechococcus* (marine cyanobacteria) spp. with rRNA-targeted peptide nucleic acid probes. *Applied and Environmental Microbiology*, *66*(1), 284-289.

Worden, A. Z., Cuvelier, M. L., & Bartlett, D. H. (2006). In-depth analyses of marine microbial community genomics. *Trends in Microbiology*, *14*(8), 331-336.

Worden, A. Z., Lee, J., Mock, T., Rouzé, P., Simmons, M. P., Aerts, A. L., Allen, A. E., et al. (2009). Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*, *324*(5924), 268-272.

Worden, A. Z., Nolan, J. K., & Palenik, B. (2004). Assessing the dynamics and ecology of marine picophytoplankton: the importance of the eukaryotic component. *Limnology and Oceanography*, *49*(1), 168-179.

Worden A. Z. & Not F. (2008). Ecology and diversity of picoeukaryotes. In: Kirchman D. L., Editor. Microbial ecology of the ocean. 2nd Edition. Wiley-Liss, New York. pp. 159–196.

Wright, S. W., Ishikawa, A., Marchant, H., Davidson, A., van den Enden, R., & Nash, G. (2009). Composition and significance of picophytoplankton in Antarctic waters. *Polar Biology*, *32*(5), 797-808.

Wright, S. W., Jeffrey, S. W., & Mantoura, R. F. C. (1997). In: Jeffrey, S. W., Mantoura R. F. C. & Wright, S. W. Evaluation of methods and solvents for pigment extraction. In: Phytoplankton pigments in oceanography: guidelines to modern methods. UNESCO publishing, Paris, France. pp. 261-282.

Xue, L., Zhang, Y., Zhang, T., An, L., & Wang, X. (2005). Effects of enhanced ultraviolet-B radiation on algae and cyanobacteria. *Critical Reviews in Microbiology*, *31*(2), 79-89.

Zapata, M., Jeffrey, S. W., Wright, S. W., Rodríguez, F., Garrido, J. L., & Clementson, L. (2004). Photosynthetic pigments in 37 species (65 strains) of Haptophyta: implications for oceanography and chemotaxonomy. *Marine Ecology Progress Series*, *270*, 83-102.

Zehr, J. P., Bench, S. R., Carter, B. J., Hewson, I., Niazi, F., Shi, T., Tripp, H. J., et al. (2008). Globally distributed uncultivated oceanic N_2-fixing cyanobacteria lack oxygenic photosystem II. *Science*, *322*(5904), 1110-1112.

Zhang, Z., Shrager, J., Jain, M., Chang, C., Vallon, O., & Grossman, A. R. (2004). Insights into the survival of *Chlamydomonas reinhardtii* during sulfur starvation based on microarray analysis of gene expression. *Eukaryotic Cell*, *3*(5), 1331-1348.

Zubkov, M. V., & Tarran, G. A. (2008). High bacterivory by the smallest phytoplankton in the North Atlantic Ocean. *Nature*, *455*(7210), 224-226.