UNIVERSITY OF CALIFORNIA

SANTA CRUZ

Quantitative biogeography of distinct clades of picoplanktonic marine green algae establishes that *Bathycoccus* ecotypes coexist more frequently than *Ostreococcus* ecotypes

A thesis submitted in partial satisfaction

of the requirements for the degree of

MASTER OF SCIENCE

in

OCEAN SCIENCES

by

Alexander Joseph Limardo

December 2016

The Thesis of Alexander Joseph Limardo
is approved:

_____

Professor Marilou Sison-Mangus

_____

Professor Alexandra Z. Worden

_____

Professor Jonathan Zehr

_____

Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures & Tables

Acknowledgements

First, I want to thank my parents, Lynn and Ken Limardo. Obviously I wouldn't exist without them, but they also helped instill in me the value of education, and as a child they nurtured my scientific curiosity. Also, none of this would have been possible without the love and support of my significant other, Amy Klebesadel, who has stood by my side throughout the good times as well as the bad.

I wanted to say a big THANK YOU to all Worden Lab members (past and present) who have helped me tremendously in so many ways these last 4 years (5 years including my time in the lab prior to becoming a grad student) as well as many MBARI staff members. A special thanks to those of you that sacrificed your time to give constructive feedback on my thesis presentation and on the thesis text itself. I feel fortunate to be a part of such a talented lab group with people who genuinely care about each other, and who are willing to help fellow lab-mates succeed in achieving their goals.

Additional thanks to the amazing UCSC Ocean Sciences Department, my thesis reading committee (Marilou Sison-Mangus, Alex Worden and Jon Zehr), and all of my Santa Cruz and MBARI friends. Last, but certainly not least, I wanted to give a special thanks to my advisor Alex Worden for all the support and guidance she has provided these last few years. During my time with Alex, she has provided me with invaluable experiences, both in my professional as well as personal life, and these experiences will continue to influence my life for many years into the future.

Abstract

Quantitative biogeography of distinct clades of picoplanktonic marine green algae establishes that *Bathycoccus* ecotypes coexist more frequently than *Ostreococcus* ecotypes

by

Alexander Joseph Limardo

We describe the biogeography of two clades of *Bathycoccus*, a picoplanktonic group of marine green algae, that are phylogenetically distinct based on analysis of the ITS1/5.8S/ITS2 region of the nuclear rRNA operon, but identical across the 18S rRNA gene. QPCR assays were developed against the ITS1 to allow quantification of these taxa (clades BI and BII). These were used alongside prior 18S rRNA primer-probes for two *Ostreococcus* clades that have previously been classified as different ecotypes (clades OI and OII) to analyze 266 photic zone samples from the tropical Atlantic and North Pacific Oceans. Clades BI and OI were observed in mesotrophic locations, whereas BII and OII were found in oligotrophic regions. *Bathycoccus* ecotypes co-occurred more often than *Ostreococcus* ecotypes. In contrast to *Ostreococcus* ecotypes, the adaptive differences between *Bathycoccus* clades have yet to be resolved and both *Bathycoccus* ecotypes appear to be more flexible in their niche(s) than *Ostreococcus*.

CHAPTER 1: Background & Introduction

Marine phytoplankton are responsible for approximately half of global net primary production, and comprise the base of the marine food web (Field et al 1998). By utilizing energy from sunlight to fix CO_2 into organic carbon molecules, the growth of phytoplankton biomass has a strong impact on the entire marine ecosystem as well as the physico-chemical conditions of the surrounding environment (Redfield 1958). Phytoplankton growth and mortality rates are controlled by numerous physical, chemical and biological factors. Understanding how these complex environmental parameters interact to determine phytoplankton community structure and biomass production is crucial to improving models of biogeochemical cycling and trophic interactions, and thus our ability to predict future changes to those systems.

Oceanographers conventionally classify plankton based on size fractions retrieved from sequential filtration through different pore sizes. The smallest size class, the picophytoplankton (<3 μm cell diameter) are widely distributed in all aquatic habitats, and comprise a diverse group of organisms, including cyanobacteria and the 'picoeukaryotes' (Johnson and Sieburth 1982, Waterbury et al 1979). Due to their small cell size which has a high surface area to volume ratio, picophytoplankton are effective competitors for nutrients, therefore this group dominates primary production in marine habitats where nutrients are scarce (Chisholm 1992, Raven 1998). However, the small size of picophytoplankton and lack of distinguishable morphological features presents major obstacles for researchers attempting to study

these organisms, therefore they were largely overlooked until advances such as flow cytometry, epifluorescence microcopy (combined with *in-situ* hybridization methods) and nucleic acid sequencing enabled more detailed investigations.

Flow cytometry, a tool utilized frequently to study picophytoplankton, allows oceanographers to discriminate, sort and quantify cyanobacteria (*Synechococcus* and *Prochlorococcus*) and photosynthetic picoeukaryotes (PPE) based on optical properties such as fluorescence of photosynthetic pigments and cell size-dependent light scattering (Olson et al 1991). Using flow cytometry to measure cell abundances, researchers began to recognize that picophytoplankton account for a significant proportion of photosynthetic biomass in many regions of the world's oceans (Campbell et al 1994, Li 1994). While cyanobacteria were often numerically more abundant than PPE, studies incorporating biomass and growth estimates found that PPE can dominate primary productivity due to their larger cell volumes and relatively high growth rates (Li 1994, Worden et al 2004). However, a significant drawback of flow cytometry is limited taxonomic resolution of picoeukaryotes, where cell abundances of many different species are often summed together and reported as a single group. When DNA sequencing techniques were applied to ocean samples, the immense genetic diversity and complexity of the PPE component was revealed, and several novel branches were added to the eukaryotic tree of life (Diez et al 2001, Lopez-Garcia et al 2001, Moon-van der Staay et al 2001). Efforts to understand the ecological roles played by different PPE taxa were confounded by massive biodiversity and a lack of quantitative observations, although certain PPE taxa

appeared more frequently in environmental clone libraries or culture collections than other taxa, suggesting higher relative abundances and ecological significance.

A number of environmental surveys targeting picoeukaryotes have recognized that the Mamiellophyceae (i.e. Class II prasinophytes) are a widespread group in many diverse ecosystems, and contain taxa with seemingly cosmopolitan distributions. The genera *Bathycoccus*, *Ostreococcus* and *Micromonas* are most often observed, and can form sizeable blooms under certain conditions (Clayton et al 2016, Collado-Fabbri et al 2011, Countway and Caron 2006, Demir-Hilton et al 2011, Not et al 2004). Phylogenetic and morphological analyses revealed that Mamiellophyceae form a basal branch on the green lineage leading to land plants, therefore it has been suggested that these phytoplankton may resemble the green algal ancestor (Bhattacharya and Medlin 1998, Lewis and McCourt 2004). However, the extremely small cell size of some Mamiellophyceae taxa appears to be a derived characteristic, resulting from extensive reductive evolution and cell miniaturization that occurred independently in at least two lineages, the Mamiellaceae (containing *Micromonas*) and the Bathycoccaceae (containing *Bathycoccus* and *Ostreococcus*; Marin 2014). The smallest of these genera, *Ostreococcus*, contains the smallest known free-living eukaryote and has often been considered the "bare limit" of eukaryotic life, possessing only the minimum essentials necessary for a photoautotroph to survive in the marine environment (Derelle et al 2006, Marin 2014).

The first quantitative environmental observations to target these three genera utilized fluorescent in-situ hybridization probes with tyramide signal amplification

(FISH-TSA) to demonstrate that Mamiellophyceae, specifically *Micromonas*, dominated primary productivity at a coastal site in the English Channel (Not et al 2004). Alternatively, quantitative polymerase chain reaction (qPCR) assays have also been employed to study the abundance of Mamiellophyceae in the environment (Countway and Caron 2006, Demir-Hilton et al 2011, Marie et al 2006, Simmons et al 2016, Zhu et al 2005). Researchers have had limited success discerning the biogeography of Mamiellophyceae clades and how they relate to ecological niches. The biogeography of *Micromonas* is complicated by the fact that widely distributed clades are often observed living in sympatry (Foulon et al 2008), however *Ostreococcus* clades are rarely found coexisting at the same location, and appear to represent distinct coastal and oceanic ecotypes (Demir-Hilton et al 2011). Until recently, *Bathycoccus* was considered a single species that could be found in many diverse marine habitats, however metagenomic and metatransciptomic studies have shown the existence of at least two clades with identical 18S rRNA gene sequences that appear to represent ecotypes analogous to *Ostreococcus* clades (Monier et al 2013, Simmons et al 2016, Vaulot et al 2012). While environmental analyses have identified genes belonging to the previously unrecognized *Bathycoccus* clade, so far there has not been a rigorous quantitative investigation of the biogeography of this genus.

In recognition of their ecological importance, representatives of all three genera were selected for whole genome analysis (one *Bathycoccus*, two *Micromonas* and three *Ostreococcus* strains), providing a powerful tool for researchers to explore the

physiology, evolution and ecology of this globally important group. Genome analyses revealed evidence of genome streamlining (i.e. reductive evolution) through compaction and elimination of non-essential features, which lowers cellular nutrient quotas and enables these phytoplankton to thrive in nutrient-poor habitats (Derelle et al 2006, Moreau et al 2012, Palenik et al 2007, Worden et al 2009). Since Mamiellophyceae display relatively simple cellular organization and a common evolutionary origin with land plants, *Ostreococcus* in particular has been utilized as a simplified model organism for investigations of plant cell biology, further elevating the importance of investigations into the ecological roles of this group (Corellou et al 2009, Derelle et al 2002, Robbens et al 2005). Among the 3 sequenced *Ostreococcus* strains, 68-71% of the genomes are protein coding sequences, and most genes are present in a single copy (Piganeau et al 2011). The only genome sequenced *Bathycoccus* strain displays a similar reduction in non-coding regions and gene redundancy, however *Bathycoccus* possesses more NH_4^+ transporters than any other genome sequenced Mamiellophyceae (six versus four in *Ostreococcus lucimarinus*) and has four highly expanded gene families that are putatively involved in scale biogenesis (Moreau et al 2012, Simmons et al 2016).

In this thesis, I used molecular techniques to explore how environmental factors interact to influence the biogeography of Mamiellophyceae taxa, specifically the genera *Bathycoccus* and *Ostreococcus* (family Bathycoccacceae). In Chapter 2, I designed two qPCR assays and applied them to a set of 266 samples to reveal abundances and distributions of two *Bathycoccus* clades (Simmons et al 2016) and

compared them to ecotypes of *Ostreococccus*, for which prior biogeographical information is known (Demir-Hilton et al 2011). The set of 266 DNA samples originated from a broad range of environmental conditions, mostly in the North Pacific Ocean, including 8 transect cruises and 2 time-series studies, which enabled statistical analyses to characterize environmental variables influencing the spatial and temporal population dynamics of particular taxa.

Collectively, the research presented here will advance understanding of an ecologically important group of marine photosynthetic picoeukaryotes. Conclusions drawn from this study can inform the parameterization of *in-silico* ecosystem models, which will improve our understanding of biogeochemical cycling and enable accurate predictions of ecosystems responses to environmental perturbations. Furthermore, understanding the biogeography of these phytoplankton will potentially lead to testable hypotheses concerning their physiology that could explain the observed geographic distributions. This investigation contributes valuable biogeographic data for a group where sparse information is currently available relative to larger, more distinguishable phytoplankton taxa.

References

Bhattacharya D, Medlin L (1998). Algal phylogeny and the origin of land plants. *Plant Physiol* **116:** 9-15.

Campbell L, Nolla HA, Vaulot D (1994). The importance of *Prochlorococcus* to community structure in the central North Pacific Ocean. *Limnol Oceanogr* **39:** 954-961.

Chisholm SW (1992). Phytoplankton size. In: Falkowski PG, Woodhead AD (eds). *Primary Production and Biogeochemical Cycles in the Sea*. Plenum Press: New York. pp 213-237.

Clayton S, Lin YC, Follows MJ, Worden AZ (2016). Co-existence of distinct Ostreococcus ecotypes at an oceanic front. *Limnol Oceanogr* doi: 10.1002/lno.10373

Collado-Fabbri S, Vaulot D, Ulloa O (2011). Structure and seasonal dynamics of the eukaryotic picophytoplankton community in a wind-driven coastal upwelling ecosystem. *Limnol Oceanogr* **56:** 2334-2346.

Corellou F, Schwartz C, Motta J-P, Sanchez F, Bouget F-Y (2009). Clocks in the green lineage: comparative functional analysis of the circadian architecture of the picoeukaryote Ostreococcus. *Plant Cell* **21:** 3436-3449.

Countway PD, Caron DA (2006). Abundance and distribution of Ostreococcus sp. in the San Pedro Channel, California, as revealed by quantitative PCR. *Appl Environ Microbiol* **72:** 2496-2506.

Demir-Hilton E, Sudek S, Cuvelier ML, Gentemann CL, Zehr JP, Worden AZ (2011). Global distribution patterns of distinct clades of the photosynthetic picoeukaryote Ostreococcus. *ISME J* **5:** 1095-1107.

Derelle E, Ferraz C, Lagoda P, Eychenie S, Cooke R, Regad F *et al* (2002). Dna libraries for sequencing the genome of Ostreococcus tauri (Chlorophyta, Prasinophyceae): The smallest free-living eukaryotic cell. *J Phycol* **38:** 1150-1156.

Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S *et al* (2006). From the Cover: Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc Natl Acad Sci U S A* **103:** 11647-11652.

Diez B, Pedros-Alio C, Massana R (2001). Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Appl Environ Microbiol* **67:** 2932-2941.

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **281:** 237-240.

Foulon E, Not F, Jalabert F, Cariou T, Massana R, Simon N (2008). Ecological niche partitioning in the picoplanktonic green alga Micromonas pusilla: evidence from environmental surveys using phylogenetic probes. *Environ Microbiol* **10:** 2433-2443.

Johnson PW, Sieburth JM (1982). In situ morphology and occurence of eukaryotic phtotrophs of bacterial size in the picoplankton of estuarine and oceanic waters. *J Phycol* **18:** 318-327.

Lewis LA, McCourt RM (2004). Green algae and the origin of land plants. *Am J Bot* **91:** 1535-1556.

Li WKW (1994). Primary production of prochlorophytes, cyanobacteria, and eukaryotic ultraplankton: Measurements from flow cytometric sorting. *Limnol Oceanogr* **39:** 169-175.

Lopez-Garcia P, Moreira D, Rodriguez-Valera F (2001). Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. *FEMS Microbiol Ecol* **36:** 193-202.

Marie D, Zhu F, Balague V, Ras J, Vaulot D (2006). Eukaryotic picoplankton communities of the Mediterranean Sea in summer assessed by molecular approaches (DGGE, TTGE, QPCR). *FEMS Microbiol Ecol* **55:** 403-415.

Marin B (2014). Reduction to essentials: Ostreococcus, the smallest free-living eukaryote. *Perspect Phycol* **1:** 7-10.

Monier A, Sudek S, Fast NM, Worden AZ (2013). Gene invasion in distant eukaryotic lineages: discovery of mutually exclusive genetic elements reveals marine biodiversity. *ISME J* **7:** 1764-1774.

Moon-van der Staay SY, DeWachter R, Vaulot D (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409:** 607-610.

Moreau H, Verhelst B, Couloux A, Derelle E, Rombauts S, Grimsley N *et al* (2012). Gene functionalities and genome structure in Bathycoccus prasinos reflect cellular specializations at the base of the green lineage. *Genome Biol* **13:** R74.

Not F, Latasa M, Marie D, Cariou T, Vaulot D, Simon N (2004). A single species, Micromonas pusilla (Prasinophyceae), dominates the eukaryotic picoplankton in the Western English Channel. *Appl Environ Microbiol* **70:** 4064-4072.

Olson R, Zettler E, Chisholm S, Dusenberry J (1991). Advances in oceanography through flow cytometry. *Particle analysis in oceanography*. Springer. pp 351-399.

Palenik B, Grimwood J, Aerts A, Rouze P, Salamov A, Putnam N *et al* (2007). The tiny eukaryote Ostreococcus provides genomic insights into the paradox of plankton speciation. *Proc Natl Acad Sci U S A* **104:** 7705-7710.

Piganeau G, Grimsley N, Moreau H (2011). Genome diversity in the smallest marine photosynthetic eukaryotes. *Res microbiol* **162:** 570-577.

Raven JA (1998). Small is beautiful: the picophytoplankton. *Funct Ecol* **12:** 503-513.

Redfield AC (1958). The biological control of chemical factors in the environment. *Amer Sci* **46**.

Robbens S, Khadaroo B, Camasses A, Derelle E, Ferraz C, Inze D *et al* (2005). Genome-wide analysis of core cell cycle genes in the unicellular green alga Ostreococcus tauri. *Mol Bio Evol* **22:** 589-597.

Simmons MP, Sudek S, Monier A, Limardo AJ, Jimenez V, Perle CR *et al* (2016). Abundance and biogeography of picoprasinophyte ecotypes and other phytoplankton in the Eastern North Pacific Ocean. *Appl Environ Microbiol* **82:** 1693-1705.

Vaulot D, Lepere C, Toulza E, De la Iglesia R, Poulain J, Gaboyer F *et al* (2012). Metagenomes of the picoalga Bathycoccus from the Chile coastal upwelling. *PLoS One* **7:** e39648.

Waterbury JB, Watson SW, Guillard RRL, Brand LE (1979). Wide-spread occurence of a unicellular cyanobacterium *Synechococcus*. *Nature* **277:** 293-294.

Worden AZ, Nolan JK, Palenik B (2004). Assessing the dynamics and ecology of marine picophytoplankton: The importance of the eukaryotic component. *Limnol Oceanogr* **49:** 168-179.

Worden AZ, Lee JH, Mock T, Rouze P, Simmons MP, Aerts AL *et al* (2009). Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes Micromonas. *Science* **324:** 268-272.

Zhu F, Massana R, Not F, Marie D, Vaulot D (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol* **52:** 79-92.

CHAPTER 2: Quantitative biogeography of distinct clades of picoplanktonic marine green algae establishes that *Bathycoccus* ecotypes coexist more frequently than *Ostreococcus* ecotypes

Summary

The Mamiellophyceae are a globally distributed class of prasinophyte algae. Several Mamiellophyceae genera are in the picoplankton size class (<2 μm diameter) and have been shown to reach high cell abundances in mesotrophic marine environments. We describe the biogeography of two clades of *Bathycoccus*, a picoplanktonic member of the Mamiellophyceae, that are phylogenetically distinct based on analysis of the ITS1/5.8S/ITS2 of the nuclear ribosomal RNA operon, but identical across the 18S rRNA gene. QPCR primer-probe sets were developed against the ITS1 to allow first ever quantification of these taxa, referred to here as clades BI and BII. These were used alongside prior 18S rRNA primer-probes for two *Ostreococcus* clades that have greater phylogenetic distance and genomic divergence, and have previously been classified as different ecotypes (OI, or Clade A and OII, or Clade B). Since *Bathycoccus* and *Ostreococcus* have two copies of the rRNA operon, values from these primer-probe sets are assumed to represent double the cell abundance. Two-hundred-sixty-six photic zone samples from the tropical Atlantic and multiple North Pacific Ocean cruises were analyzed. Among the major transects or time-series, the

Ostreococcus maxima was in the coastal upwelling of the eastern North Pacific for OI (36,713 ± 1,485 18S copies ml^{-1}) and at the intersection of the Kuroshio and Oyashio Currents for OII (50,189 ± 561 18S copies ml^{-1}). For *Bathycoccus*, the maxima were observed along Line P for BI (10,667 ± 1,299 ITS copies ml^{-1}) and in the tropical Atlantic for BII (4,125 ± 339 ITS copies ml^{-1}). All four clades were found in multiple surface samples as well as deeper in the water column. The two *Ostreococcus* ecotypes only co-occurred in frontal regions where oligotrophic waters mingled with more coastal waters. In warm, oligotrophic waters BII and OII were detected but not BI and OI. However, the BII clade extended into cooler more nutrient rich waters than OII. Hence, *Bathycoccus* BI and BII overlapped in 85 of 266 samples. Depth, phosphate and salinity had significant positive relationships with BII abundance, and with the abundance ratio of these two clades in samples where both were detected. Moreover, BII was present in samples where phosphate was below detection (<10 nM) and temperatures up to 29^{o}C, whereas BI was only found in samples with phosphate >20 nM and temperature ≤25^{o}C over the whole sample set. *Ostreococcus* OI and OII overlapped in fewer samples (42), 76% of which were from two cruises in frontal zones. Notably, none of these taxa are discriminated by the 18S rRNA gene V9 region, which severely limits interpretation of plankton distributions and trajectories in future warming scenarios that are based on this commonly used marker. Collectively, our studies establish that there are distinct ecotypes of *Bathycoccus*. However, unlike the *Ostreococcus* clades which can clearly be defined as mesotrophic (OI) and oligotrophic (OII), the adaptive differences of the

Bathycoccus ecotypes have yet to be resolved and both *Bathycoccus* ecotypes appear to be more flexible in their niche(s) than the *Ostreococcus* ecotypes.

Introduction

Prasinophytes are green algae that are present in many marine environments. Members of the class Mamiellophyceae appear to be particularly widespread and include several picoplanktonic (<2 µm diameter) genera. 18S rRNA gene clone libraries from the picoplankton size class show numerous sequences from the Mamiellophyceae genera *Bathycoccus*, *Micromonas*, and *Ostreococcus* (Worden and Not 2008) as do 18S rRNA gene amplicon sequences from 47 sites at different locations in the global ocean (de Vargas et al 2015). However, because of the non-quantitative nature of these studies, and limited phylogenetic resolution of markers such as the 18S rRNA V9 region, see e.g. (Monier et al 2016), relatively little is known about the biogeography of the distinct clades that comprise these genera or their numerical contributions. *Ostreococcus* has been shown to dominate the picoeukaryotic phytoplankton community numerically off coastal Chile (Collado-Fabbri et al 2011) and has been reported in high numbers at several other mesotrophic regions (Clayton et al 2016, Countway and Caron 2006, Simmons et al 2016). *Bathycoccus* and *Micromonas* seem to be lower in abundance in these studies, but are larger than *Ostreococcus* and hence can account for more of the picophytoplankton biomass (Collado-Fabbri et al 2011). In the more oligotrophic Sargasso Sea all three of these genera are more abundant during the spring bloom and decline to low or undetectable numbers when the water column becomes strongly stratified (Treusch et al 2012). A summary of the literature to date indicates that *Bathycoccus* and

Micromonas are present in surface waters from low to high latitudes, including polar seas, whereas *Ostreococcus* appears to be restricted to temperate and tropical waters (Guillou et al 2004, Not et al 2005, Simmons et al 2015, Vaulot et al 2008).

Understanding the dynamics and contributions of *Bathycoccus*, *Micromonas*, and *Ostreococcus* is complicated by the fact that each genus contains several species and / or phylogenetically distinct clades that as yet cannot be distinguished based on morphological characteristics. Genome sequencing and evolutionary analyses have underscored the differences between these distinct clades see e.g., (Moreau et al 2012, Palenik et al 2007, Rodriguez et al 2005, Simmons et al 2015, Simmons et al 2016, van Baren et al 2016, Worden et al 2009). The *Micromonas* genus has seven distinct clades, six of which are cultured, and exhibit the most genetic diversity, including two formally described species, *Micromonas commoda* and *Micromonas pusilla* (van Baren et al 2016). *Ostreococcus* has four well supported clades based on the 18S rRNA gene and Intergenic Transcribed Spacer (ITS) all of which contain cultured representatives, specifically, *Ostreococcus lucimarinus* (*Ostreococcus* clade A), *Ostreococcus* sp. RCC809 (clade B) *Ostreococcus tauri* (clade C) and *Ostreococcus mediterraneus* (clade D) (Guillou et al 2004, Rodriguez et al 2005, Subirana et al 2013). *Ostreococcus* clades A and B correspond to the OI and OII ecotypes, respectively, for which quantitative polymerase chain reaction (qPCR) primer-probe sets have been developed and implemented (Demir-Hilton et al 2011). The latter study concluded that OI and OII occupy different niches. *Ostreococcus* clade OI was observed in cooler, coastal locations while clade OII was observed in warmer

offshore waters throughout the water column, except during periods of stratification when it was localized to the deep chlorophyll maximum (DCM). Additionally, the OI and OII clades co-occurred at only two stations near the continental slope where physical mixing occurs between coastal and oligotrophic waters, leading to the conclusion that the clades represent ecotypes adapted to mesotrophic (OI) and oligotrophic (OII) conditions. More recently, a study of the dynamic Kuroshio Front used qPCR to confirm that OI and OII co-exist when in frontal systems (Clayton et al 2016). Interestingly, abundances of both *Ostreococcus* ecotypes reached their highest values in locations where mixing of the two water masses appeared to be strongest, possibly leading to an injection of nutrients to the surface layer.

In contrast to *Micromonas* and *Ostreococcus*, *Bathycoccus* appears to have lower genetic divergence, at both the individual marker and genomic levels (Guillou et al 2004, Simmons et al 2016). All sequenced isolates of *Bathycoccus* have identical 18S rRNA gene sequences and have been considered a single species, *Bathycoccus prasinos*. But targeted metagenomic studies demonstrated diversity in ITS2-5.8S sequences and other gene variants were recovered from a *Bathycoccus* population sorted by flow cytometry (Monier et al 2012, Monier et al 2013, Vaulot et al 2012). For example, based on a limited sample set, Monier and colleagues noted that the variant sequences appeared to occur only in oligotrophic environments whereas *B. prasinos* sequences were found in coastal settings. These studies led to the hypothesis at least two *Bathycoccus* ecotypes exist (Monier et al 2013, Vaulot et al 2012). A follow-up study used clusters of single-copy homologous genes (i.e. ‘ecomarkers’) in

metagenomics and metatranscriptomic data to quantitate nucleotide differences between the putative *Bathycoccus* ecotypes as well as the known *Ostreococcus* clades (Simmons et al 2016). This showed that the two *Bathycoccus* clades had 82 ± 6% nucleotide identity across homologs, whereas *Ostreococcus* clade OI and OII homologs displayed 75 ± 8% nucleotide identity. The latter study also demonstrated similar patterns in OI and OII distributions (largely non-overlapping; by qPCR) as a prior report on the same eastern North Pacific region and showed *Bathycoccus* to range up to 21,000 18S gene copies mL^{-1} but did not resolve the two potential *Bathycoccus* ecotypes, BI and BII.

Bathycoccus and *Ostreococcus*, the focus of our study, are evolutionarily closer to each other than to *Micromonas* and are both non-motile, unlike most prasinophytes. We set out to test whether the two proposed BI and BII groups exhibit abundance distributions like *Ostreococcus*, where the OI and OII clades rarely overlap except in dynamic frontal regions (Clayton et al 2016, Demir-Hilton et al 2011, Simmons et al 2016). To this end, we developed qPCR assays that resolve *Bathycoccus* types that form two bootstrap supported clades based on the ITS but have identical 18S rRNA gene sequences. ITS1/5.8S/ITS2 clone libraries were constructed from four sites and three cultured strains were sequenced to increase the available data for this marker. For comparison, we also quantified OI and OII using previously developed qPCR primer-probe pairs (Demir-Hilton et al 2011) in the same sample sets or used published *Ostreococcus* data for a subset of these (Clayton et al 2016, Demir-Hilton et al 2011, Simmons et al 2016). The samples came from several regions of the North

Pacific Ocean as well as the tropical Atlantic, and included two time-series sites, Station ALOHA near Hawaii and Station M1 in Monterey Bay. The abundance and distributions of these four Mamiellophyceae groups were analyzed in the context of available environmental data. The data reveal that the two *Bathycoccus* clades do represent different ecotypes with distinct biogeographies. However the BI and BII clades overlap more frequently than the two previously described *Ostreococcus* ecotypes.

Materials and Methods

Sample collection

Environmental samples were collected from locations in the North Pacific Subtropical Gyre (HOT station ALOHA), sub-Arctic Pacific (Line P), California Current System (Line 67), Kuroshio Front, tropical Atlantic and along the western coast of North America from tropical to temperate latitudes. The entire dataset included 266 samples, 78 of which come from the Hawaii Ocean Time-series (HOT) station ALOHA and 22 of which come from the Monterey Bay Time-series (MBTS) station M1. Niskin bottles on a rosette were used to collect the majority of seawater samples (Sea-Bird Electronics, Bellevue, WA, USA). Temperature, salinity, pressure (depth), sigma T (density) were collected using Conductivity, Temperature and Depth (CTD) and chlorophyll *a* fluorescence (Sea-Bird Electronics) sensors mounted on the rosette. DNA samples were collected by filtering 500-2000 mL of seawater onto 0.2 µm pore-size Supor filters (Pall Scientific, Port Washington, NY, USA) and flash-frozen in liquid N_2 or -80°C as described in Demir-Hilton et al (2011). The exception was Line P surface samples which were collected using a pump (McLane Research Laboratories, East Falmouth, MA USA).

Inorganic nutrients NO_3^-, NO_2^-, PO_4^{3-}, $Si(OH)_4$ and chlorophyll *a* samples were collected and analyzed according to previously described procedures (Pennington and Chavez 2000), or for NH_4^+ samples as in Santoro et al (2013) or Holmes et al (1999). Environmental metadata for Hawaii Ocean Time-series (HOT) samples were

retrieved from the open-source database (http://hahana.soest.hawaii.edu/hot/hot-dogs/).

Location	Cruise	Station(s)	Start or station		Site farthest from start		Dates
			Lat	Lon	Lat	Lon	
North Pacific Subtropical Gyre	HOT	ALOHA	22.75	−158.00	NA	NA	06/11–08/12
Monterey Bay	MBTS	M1	36.75	−122.00	NA	NA	03/14–11/14
Eastern N. Pacific (Line 67)	WFAD09*	9	36.80	−121.85	33.29	−129.43	10/09
Eastern N. Pacific	CANON11	3	36.13	−123.49	33.95	−128.05	09/11
Eastern N. Pacific	CANON10	2	35.92	−122.90	35.79	−122.74	09/10
Eastern N. Pacific	MV1405	3	38.24	−126.62	35.93	−121.73	07/14
Monterey Bay to Gulf of California	GOC12	9	36.70	−122.38	21.00	−110.00	02/12
N. east sub-Arctic Pacific (Line P)	GeoMICS	4	48.58	−125.50	48.82	−128.67	05/12
Kuroshio Current Frontal Zone	NT09-18*	25	36.61	143.48	35.17	145.5	10/09
Tropical Atlantic	SJ0609*	8	11.67	−51.49	1.83	−38.49	06/06–07/06

Table 1. Cruises analyzed herein. Those for which *Ostreococcus* results have previously been reported (Clayton et al 2016, Demir-Hilton et al 2011, Simmons et al 2016) are denoted (*). Note that 10 HOT cruises to Station ALOHA were analyzed and 11 MBTS cruises to Station M1.

Culturing

Bathycoccus CCMP1898 was acquired from National Center for Marine Algae and Microbiota (NCMA) at Bigelow Laboratory (East Boothbay, ME, USA), and *Bathycoccus* RCC715, and RCC716 were acquired from the Roscoff Culture Collection (Roscoff, France) and grown in L1 media (Guillard and Hargraves 1993) at 21°C. Cells were grown in semi-continuous batch culture on a 14:10 hr light/dark cycle at an irradiance of 25 μmol m^{-2} s^{-1} for RCC715 and RCC716, and at 85 μmol m^{-2} s^{-1} for CCMP1898. Cell growth was monitored on an Accuri C6 flow cytometer (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). For DNA extracts and subsequent ITS sequencing, cultures were filtered onto 0.2 μm pore-sized Supor filters (Pall Scientific) and flash frozen in liquid N_2. Strains RCC715 and RCC716 were used for ITS sequencing, and CCMP1898 and RCC716 were used to test qPCR assays.

DNA extraction and clone libraries

DNA was extracted using the DNeasy Plant kit (Qiagen, Hilden, Germany) and a modified version of the extraction protocol, which included a mechanical lysis step in addition to a freeze-thaw step to ensure disruption of the cell wall (Moisander et al

2008). DNA was eluted in 50 µl of TE buffer pH 8. 1:4 dilutions were prepared and aliquoted before storage at -80°C, in order to minimize freeze-thaw cycles and subsequent degradation of DNA.

The primers used for ITS clone library construction consisted of a general eukaryote reverse primer that anneals to a conserved region the beginning of the 28S rRNA gene (ITS055R: 5'-CTCCTTGGTCCGTGTTTCAAGACGGG-3'; Marin et al 1998) and a forward primer (MAM18S1F: 5'-CGAAAGGTCTRGGTAATCTCC-3') designed here to anneal to the 18S rRNA gene of *Bathycoccus, Micromonas,* and *Ostreococcus* and therefore recover the ITS1, 5.8S and ITS2 along with flanking 18S and 28S gene regions. PCR reactions were performed using HotStar Taq polymerase (Qiagen) and thermal cycling conditions were as follows: 95°C for 15 min, followed by 30 cycles of 95°C for 30s, 60°C for 30s, and 72°C for 1.5 min, and final extension was 72°C for 7 min. The amplicons were then cloned using the pCR2.1-TOPO kit (Thermo Fisher Scientific, Waltham, MA USA). Plasmids were isolated using a Qiagen Spin Miniprep Kit (Qiagen) and amplicons were sequenced using the primers M13F and M13R on the Applied Biosystems 3500xL Genetic Analyzer using Big Dye Terminator v3.1 sequencing kit (Thermo Fisher Scientific).

Phylogenetic tree construction

Clone library sequences as well as ITS sequences from targeted metagenomes (Chile 1 & 2; Tropical Atlantic) from prior studies were included (Monier et al 2013, Vaulot et al 2012) as were other representative Mamiellophyceae sequences retrieved

from GenBank. A total of 47 sequences covering the entire ITS1, 5.8S and ITS2 regions were aligned using MAFFT Version 7 (Katoh and Standley 2013) with default settings. Gap positions were discarded. A maximum likelihood (ML) phylogenetic tree was constructed based on 686 nucleotide positions using MEGA6 (Tamura et al 2013) applying a general time reversible model with invariant sites and a gamma distribution of rates (GTR+I+G). Bootstrap values were calculated from 1000 replications. Only representative sequences from the Kuroshio Front, tropical Pacific and tropical Atlantic environmental clone libraries, and Monterey Bay were included for sequences within a clone library that had >99% nucleotide identity.

Quantitative real-time polymerase chain reaction

Two qPCR assays were developed for two *Bathycoccus* clades (referred to here as clades BI and BII, Fig. 1, Tables 2, 3). Probe sequences were identified using Beacon Designer software (PREMIER Biosoft, Palo Alto, CA USA) and primers were designed manually based on alignments of 47 Mamiellophyceae ITS sequences including representatives of all known genera, using MEGA6 (Tamura et al 2013). The specificity of both new qPCR assays was tested on serial dilutions (10^8-10^1 gene copies well^{-1}) of plasmids bearing ITS genes from *Micromonas*, *Ostreococcus* and *Bathycoccus* either recovered from environmental clone libraries or from cultured strains (Table 3). Furthermore, extracts of *Bathycoccus* sp. CCMP1898 and RCC716 genomic DNA were tested with both primer/probe sets and non-target amplification was never observed (Fig. 2). Amplification efficiency was computed as %

efficiency=($10^{(-1/\text{slope of linear regression})}$-1)*100) over 3 technical replicates at each of 8 plasmid concentrations, and 6 technical replicates at the lowest concentration of plasmid (range as above), with a total of 9 plasmid concentrations tested.

Environmental DNA samples were enumerated by absolute quantification using the standard curve method. Samples were measured in triplicate, along with a test for inhibition and no-template controls. Reactions were performed in 25 µl volumes on an Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA USA) and were comprised of Taqman Universal PCR master mix (Thermo Fisher Scientific), 900 nM of forward and reverse primers, 250 nM of probe for BI, OI and OII assays, and 150 nM of probe for BII assay. Nuclease-free water and TE buffer were used as no-template controls. QPCR reactions were as follows: initial 10 min at 95°C, followed by 45 cycles at 95°C (15 sec), then 60°C (1 min), and fluorescence data were collected during the annealing/extension phase. Baseline values were calculated using the AB7500 software, and threshold value was set to default value of 0.2. Standard curves consisted of a known DNA template in a plasmid vector (pCR 2.1-TOPO). Template DNA for *Bathycoccus* assays consisted of environmental ITS sequences obtained in this study, and for *Ostreococcus* assays consisted of 18S rRNA sequences obtained from cultured representatives as described in Demir-Hilton et al (2011). DNA concentration of standards were quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific), and standard curves were prepared by ten-fold serial dilution (10^8-10^1 gene copies $well^{-1}$), aliquoted into single-use tubes (to minimize freeze-thaw cycles) and stored at -20°C. The detection limit

for quantification of our assays was 10 template copies well^{-1}. Abundances below the detection limit were retained as detected but not quantifiable, and used to define presence versus below detection and/or absent. Sample inhibition was tested by spiking in 10^5 template copies well^{-1} as described previously (Demir-Hilton et al 2011). DNA samples were diluted between 1:20 and 1:200 to prevent inhibition, although a 1:40 dilution was typically sufficient. 18S rRNA gene counts for *Ostreococcus* clade OI and OII have been published for cruises SJ0609 (Demir-Hilton et al 2011), WFAD09 (Simmons et al 2016), and the Kuroshio Front transect (Clayton et al 2016).

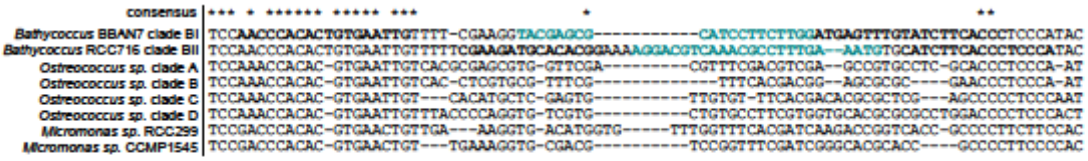


Figure 1. Alignment of Mamiellophyceae partial ITS1 sequences. QPCR primer regions targeted by assays designed in this study for *Bathycoccus* clades BI and BII are highlighted in black, and probe sequences are in blue.

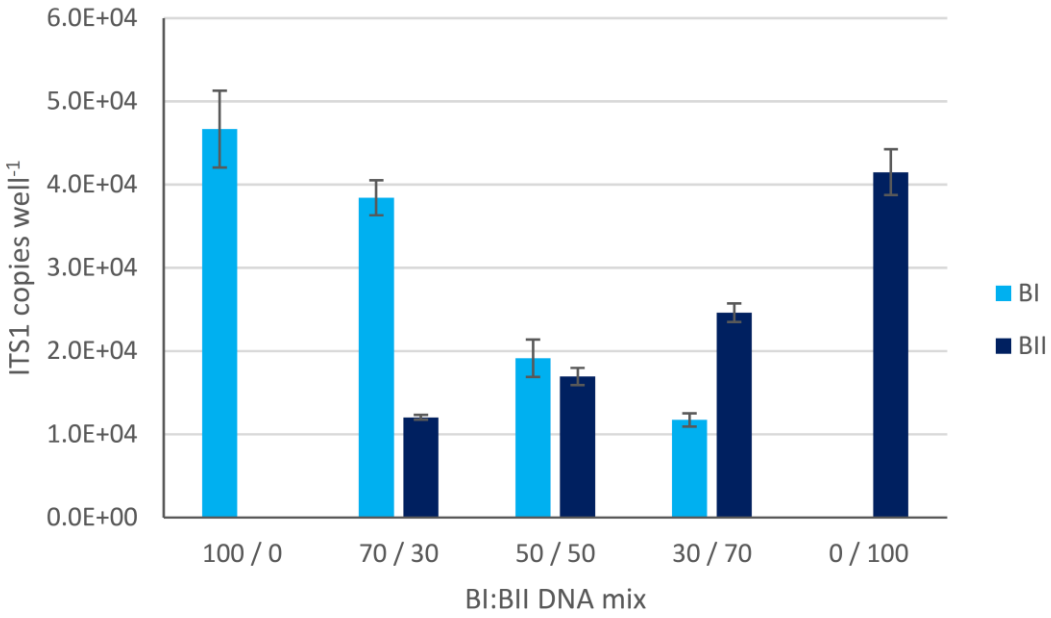


Figure 2. Results of *Bathycoccus* qPCR assays tested on mixtures of genomic DNA from *Bathycoccus* CCMP1898 (clade BI) and RCC716 (clade BII).

Targeted clade	Name	Sequence (5' - 3')
BI	B1.ITS1.3F	AACCCACACTGTGAATTG
	B1.ITS1.3R	GGGTGAAGATACAAACTCAT
	B1.ITS1.3probe	TACGAGCGCATCCTTCTTGG
BII	B2.ITS1.3F	CGAAGATGCACACGG
	B2.ITS1.3R	TGGGAGGGTGAAGATG
	B2.ITS1.3probe	CATTTCAAAGGCGTTTGACGTCCT

Table 2. Primer and probe nucleotide sequences for qPCR assays developed to quantify *Bathycoccus* clades BI and BII.

Organism source material						primer/probe assay	
Class	**Genus**	**Species**	**Strain/Clone**	**Clade**	**DNA**	**B1.ITS1.3**	**B2.ITS1.3**
Mamiellophyceae	*Bathycoccus*	*prasinos*	CCMP1898	BI	C	Y	UN
Mamiellophyceae	*Bathycoccus*	*prasinos*	HM12	BI	E	Y	UN
Mamiellophyceae	*Bathycoccus*	*sp.*	RCC716	BII	C	UN	Y
Mamiellophyceae	*Bathycoccus*	*sp.*	TAM10	BII	E	UN	Y
Mamiellophyceae	*Ostreococcus*	*lucimarinus*	CCMP2972	A.OI	C	UN	UN
Mamiellophyceae	*Ostreococcus*	*sp.*	RCC809	B.OII	C	UN	UN
Mamiellophyceae	*Ostreococcus*	*sp.*	TAM3	B.OII	E	UN	UN
Mamiellophyceae	*Micromonas*	*sp.*	KM4	A.II	E	UN	UN
Mamiellophyceae	*Micromonas*	*sp.*	HM4	C.I	E	UN	UN
Mamiellophyceae	*Micromonas*	*sp.*	HM6	E1	E	UN	UN

Table 3. Specificity tests of qPCR assays targeting *Bathycoccus* clades using genetic material derived from cultures or environmental clones. C - DNA sourced from cultured strains; E - environmental clone ITS sequences; Y - detected by qPCR; UN - undetected.

Statistical analyses of qPCR and environmental data

Matrices for statistical analyses were assembled containing clade qPCR abundance data and the following environmental variables: temperature, salinity, density (σ_T), depth, NO_3^-, NO_2^-, NH_4^+, PO_4^{3-}, and chlorophyll *a*. A matrix was assembled including all available samples, as well as matrices for individual transect cruises and time-series studies in order to evaluate distinct environmental regimes. Non-metric multidimensional Scaling (NMDS) analysis was performed on a dissimilarity matrix containing 213 samples and 6 variables (temperature, salinity, depth, NO_3^-, PO_4^{3-}, N:P ratio). Data was z-scored to standardize units before using Euclidean distances to

calculate the dissimilarity matrix. NMDS coordinates were then used in a k-means cluster analysis (based on square Euclidean distance, 50 replications) to divide samples into three groups. Wilcoxon rank-sum tests, Spearman's rank correlation tests, Kruskal–Wallis tests, Kolmogorov–Smirnov tests, Lilliefors tests, NMDS and k-means clustering were performed using MATLAB (The MathWorks Inc., Natick, MA USA). Abundance ratios for WFAD09 and GOC12 qPCR data were constructed by dividing abundances (e.g. clade BI/clade BII). For temperature and salinity distributions of *Bathycoccus* and *Ostreococcus* clades observation numbers differed for *Bathycoccus* clade BI (152 observations), clade BII (153 observations), *Ostreococcus* clade OI (131 observations), and clade OII (124 observations) and samples where the clade was detected by qPCR but not quantifiable (i.e., <10 copies $well^{-1}$) were included as observations.

Results

Bathycoccus and *Ostreococcus ITS Phylogeny*

ITS1/5.8S/ITS2 clone libraries were constructed from four samples, providing environmental *Bathycoccus* ITS sequences from the tropical Atlantic, tropical Pacific, Kuroshio Front, and Monterey Bay (Fig. 3A, Tables 1, 4). *Ostreococcus* were recovered at each of these sites except Monterey Bay. Sequences from *Micromonas* and other algae were recovered as well, but not analyzed further. Our maximum-likelihood phylogenetic analyses focused on resolving *Bathycoccus* clades (Fig. 3B) and the overall topology was similar to a prior reconstruction for Mamiellophyceae (Marin and Melkonian 2010). The exception was that we identified two statistically resolved clades (100% bootstrap support) for *Bathycoccus* as opposed to one. This is likely because the prior study only analyzed cultured taxa with deposited sequences, while here we sequenced and analyzed additional isolates (RCC715 and RCC716) from the Indian Ocean, in addition to analyzing sequences from the environmental clone libraries. Here, Monterey Bay *Bathycoccus* ITS sequences had on average 99% nucleotide identity to the genome strain, *Bathycoccus prasinos* BBAN7 (isolated from the Mediterranean near the southern coast of France). In contrast, the majority of sequences from the other three sites, as well as the two Indian Ocean isolates, belonged to a distinct clade. Environmental *Bathycoccus* clones had 95% average ITS nucleotide identity to BBAN7, except one tropical Pacific clone which was 99% identical to BBAN7. Sequence divergence between genotypes was in the ITS1 and

ITS2 regions, as 5.8S nucleotide sequences were 100% identical between all environmental clones and cultured *Bathycoccus prasinos* strains. The *Bathycoccus* clade containing known cultured isolates and the genome sequenced *Bathycoccus prasinos* strain BBAN7 is hereafter referred to as "clade BI" and the other distinct clade as "clade BII," according to the proposal of Simmons et al (2016).

Location	Abbreviation	Latitude	Longitude	Genera	Number of clones
Kuroshio Front	KM	35.50°N	145.52°E	*Bathycoccus*	2
				Ostreococcus	4
				Micromonas	4
Monterey Bay	HM	36.74°N	122.02°W	*Bathycoccus*	6
				Micromonas	5
Tropical Atlantic	TAM	12.41°N	27.25°W	*Bathycoccus*	5
				Ostreococcus	17
Tropical Pacific	G	21.00°N	110.00°W	*Bathycoccus*	6
				Ostreococcus	6

Table 4. Results from environmental clone libraries targeting Mamiellophyceae.

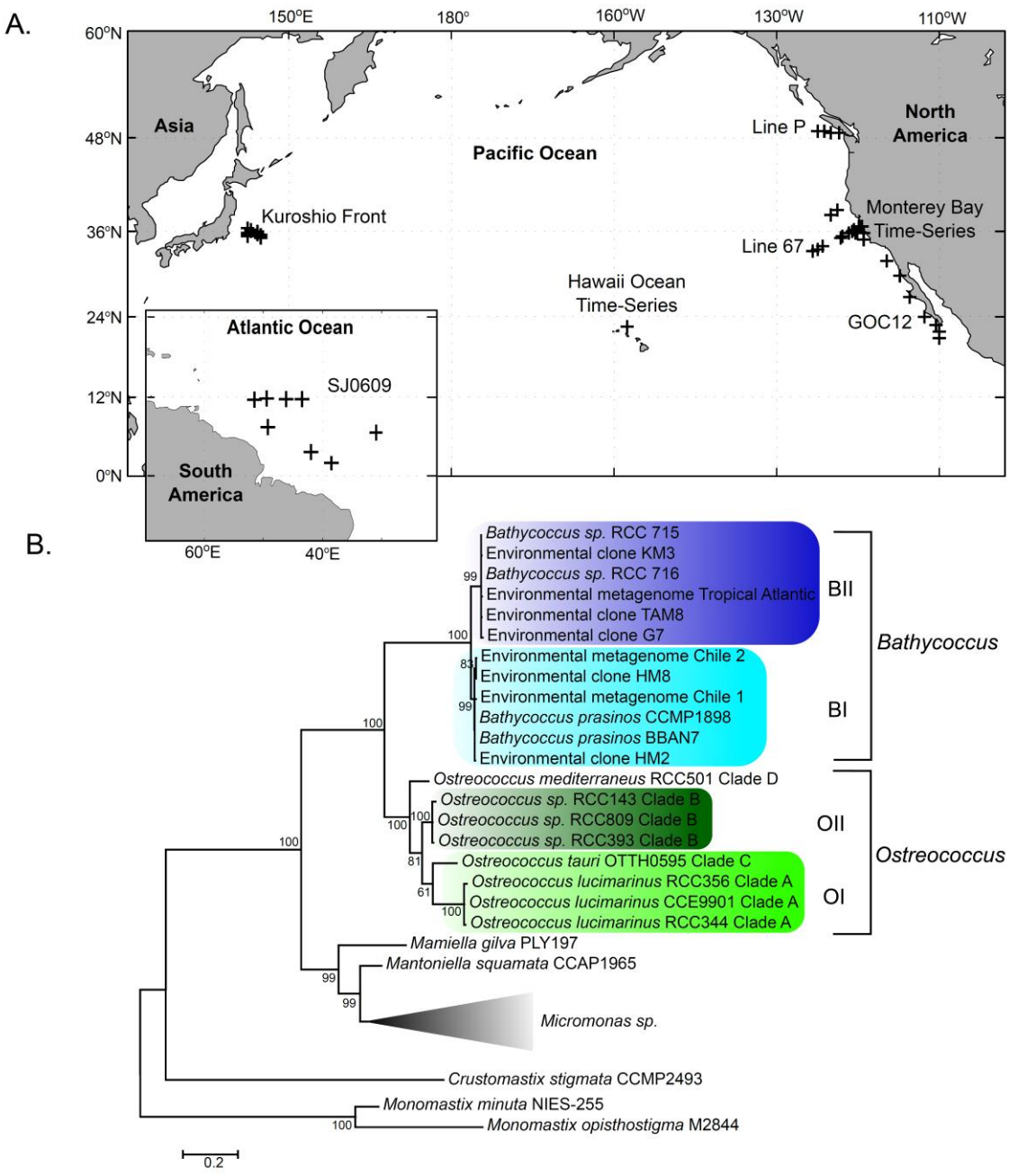


Figure 3. Delineation of two *Bathycoccus* clades using environmental samples and phylogenetic analysis of the ITS1/5.8/ITS2. **A.)** Locations sampled in this study are denoted by a black cross (+). **B.)** Maximum likelihood tree of Mamiellophyceae ITS1/5.8/ITS2 region of the nuclear genome rRNA operon, with an emphasis on *Bathycoccus* and *Ostreococcus*. Environmental *Bathycoccus* ITS clones from 4 locations sequenced in this study (bold) are included, but only one or two

representative sequences from each location were retained for this tree. Three metagenomics sequences (Chile 1 & 2; Tropical Atlantic) from two prior studies (Vaulot et al 2012, Monier et al 2012) were also included. Numbers at nodes represent bootstrap values (percent of 1000 replicates).

QPCR Assays to Distinguish Bathycoccus and *Ostreococcus Clades*

Environmental *Bathycoccus* ITS sequences and sequences from cultured isolates were used to design qPCR assays that distinguished clades BI and BII (Fig. 1, Tables 2, 3). The BI and BII primer-probe sets were used alongside published primer-probe sets for *Ostreococcus* clade OI and clade OII, which were previously established to have 90-101% (OI) and 91-99% (OII) efficiencies (Demir-Hilton et al 2011). Mean amplification efficiencies for the qPCR assays were 97.7 ± 2.1% (n=13) for the BI assay, 91.0 ± 1.3% (n=13) for the BII assay, and the published *Ostreococcus* assays showed efficiencies of 98.6 ± 5.1% (n=14, OI), and 91.2 ± 2.2% (n=15, OII) herein. However, the y-intercepts were high (41-43 for BI, OI, OII and 46-47 for BII), suggesting that the assays as implemented here may underestimate actual abundances, especially for the BII assay (Bustin 2004). Finally, the sequenced *Bathycoccus prasinos* BBAN7 and *Ostreococcus lucimarinus* CCE9901 genomes (Moreau et al 2012, Palenik et al 2007) show two copies of the rRNA operon that contains both the 18S rRNA gene and the ITS. Therefore, qPCR data shown here are presumed to represent double the cell abundance.

Temporal Dynamics of Bathycoccus and Ostreococcus Clades

We investigated temporal changes in the abundance of the *Bathycoccus* and *Ostreococcus* clades at Station M1 in Monterey Bay, where both genera had previously been enumerated (Demir-Hilton et al 2011, Simmons et al 2016), and at Station ALOHA in the North Pacific Subtropical Gyre (spanning 14 months/10 sampling dates). The Monterey Bay study spanned March to November 2014 (11 sampling dates). Surface (<5 m) temperatures were lowest (11.4°C) in early May when salinity, chlorophyll *a* and surface concentrations of nutrients (NO_3^-, NO_2^-, PO_4^{3-}, and SiO_4^{2-}) were maximal, indicating the influence of wind-driven coastal upwelling (Fig. 4, Supplemental Datafile). Thermal stratification increased during summer as surface temperature increased from 13.6°C on 10 July to 17.0°C on 12 August and peaked at 17.3°C on the September 22 sampling date. Long-term observations of the Monterey Bay show that surface temperatures rarely exceed 15°C during the late summer and autumn (Pennington and Chavez 2000), thus, our study includes a period of unusually warm water temperatures (Fig. 4) associated with an anomalously warm patch of water present in the eastern North Pacific throughout 2014 (Bond et al 2015). Additionally, NO_3^- was <1.0 µM from 12 August until the end of the study period and PO_4^{3-} decreased from a maximum (1.4 µM) on 5 May to a minimum (0.2 µM) on 29 October (Supplemental Datafile). Salinity also decreased between July and late October, reflecting shoreward movement of warmer, low salinity California Current waters.

Bathycoccus and *Ostreococcus* qPCR results showed differential temporal distributions in Monterey Bay during the study period (Fig. 4). The two depths

analyzed, surface (<5 m) and the base of the wind-mixed surface layer (~10-20 m), or chlorophyll maximum if present, trended together over time. *Bathycoccus* clade BI was detected throughout the study period and was most abundant in October and November with a peak of 8,444 ± 529 ITS copies mL^{-1} (20 m, 20 November). *Bathycoccus* clade BII was detected from August to November, with a peak of 775 ± 71 ITS copies ml^{-1} (20 m, 20 November) and its overall abundance was significantly lower than BI (Wilcoxon rank sum test, $p<0.0001$). *Ostreococcus* clade OI was also present in Monterey Bay throughout the entire study period, higher in abundance in the spring and summer, reaching a maximum on 19 June (18,033 ± 1,260 18S copies mL^{-1}). *Ostreococcus* clade OII was not detectable on 7 of the 11 sample collection dates at either depth. OII was detected between August, and November when waters were warmest, but did not exceed 50 18S copies mL^{-1}. Since clades BI and OI were present in all samples, and clades BII and OII were only detected on four of the 11 sampling dates, we compared environmental variables from Monterey Bay samples where *Bathycoccus* clade BII was present versus absent, and performed the same analysis with *Ostreococcus* clade OII (Table 5). In samples where *Bathycoccus* clade BII was detected, temperature (15.7 ± 1.0 °C) and salinity (33.31 ± 0.11) were significantly different than temperature (13.4 ± 2.0 °C) and salinity (33.49 ± 0.21) of samples where they were absent (Wilcoxon rank sum, $p<0.05$). Additionally, NH_4^+ (0.06 ± 0.08 µM), NO_3^- (1.22 ± 2.29 µM) and N:P (1.4 ± 1.9) were lower ($p<0.05$) than samples where this clade was absent (NH_4^+, 0.25 ± 0.25 µM; NO_3^-, 4.57 ± 3.26 µM; and N:P, 4.6 ± 2.6). *Ostreococcus* clade OII was also found in higher

temperature waters (15.8 ± 1.1 °C), with lower NH_4^+ (0.07 ± 0.07 µM), NO_3^- (1.20 ± 2.31 µM), and N:P (1.3 ± 1.9), however statistical differences were not observed for salinity.

The Station ALOHA samples analyzed here were collected on 10 sampling dates distributed between July 2011 and August 2012. Over this study period, strong thermal stratification was observed in the photic zone (<200 m), which weakened in the winter months (Fig. 5). Sea surface temperatures varied from a maximum of 26.3°C in September 2011 to a minimum of 23.0°C in March 2012. The DCM was located between 100 m to 150 m in depth. Average NO_3^- (0.02 ± 0.01 µM) and PO_4^{3-} (0.09 ± 0.05 µM) concentrations were persistently low at the surface (5 m, Supplemental Datafile). Neither *Bathycoccus* clade BI nor *Ostreococcus* clade OI were detected at Station ALOHA, whereas *Bathycoccus* clade BII and *Ostreococcus* clade OII were present year-round deep in the photic zone.

Bathycoccus clade BII and *Ostreococcus* clade OII were present at low abundance in ALOHA surface depths ≤50 m from November to May and otherwise below detection limits during warmer, strongly stratified periods (Fig. 5). Generally their maximum abundance in the water column followed the DCM. The highest numbers detected were comparable between these two taxa, with 1,146 ± 135 ITS copies mL^{-1} (*Bathycoccus* clade BII) and 1,067 ± 52 18S copies mL^{-1} (*Ostreococcus* clade OII). However, these maxima occurred on different sampling dates with *Bathycoccus* clade BII's peak occurring as the DCM shallowed just prior to the winter mixing period (100 m, 4 November 2011), and *Ostreococcus* clade OII's peak occurred deeper in

the water column as thermal stratification strengthened in spring (150 m, 30 May 2012). PO_4^{3-} was not significantly different between samples containing *Bathycoccus* clade BII and *Ostreococcus* OII and those without either (0.1 ± 0.04 and 0.09 ± 0.04 µM, respectively). However, NO_3^- was significantly higher in samples containing these taxa versus those where they were below detection (0.36 ± 0.54 and 0.02 ± 0.01 µM, respectively, $p<0.01$, Table 6). This suggests that their presence deep in the photic zone is driven by nutrient availability, particularly nitrogen or co-associated variables, and is in agreement with the hypothesis that primary production in the North Pacific Subtropical Gyre is limited by the availability of fixed nitrogen sources (Moore et al 2013, Tyrrell 1999).

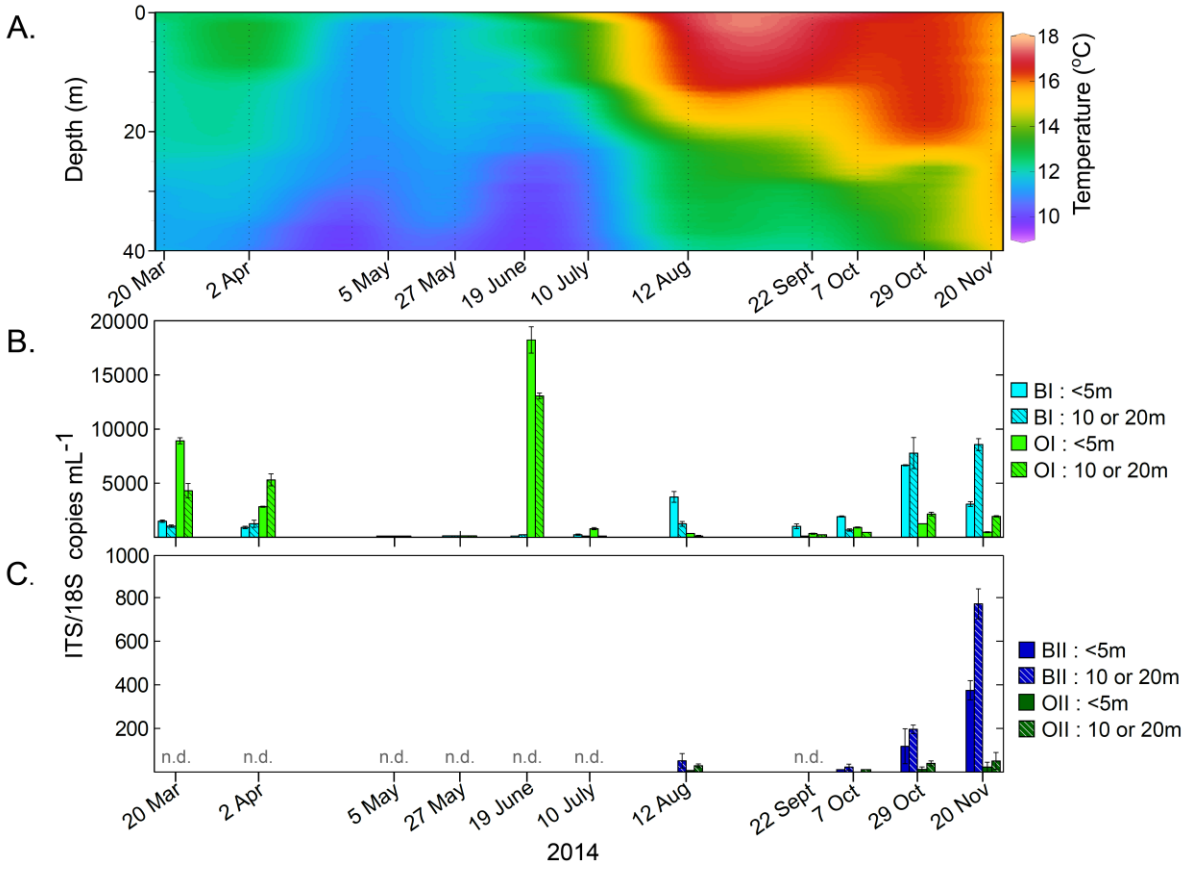


Figure 4. Dynamics of *Ostreococcus* and *Bathycoccus* clades in Monterey Bay, California. Samples were collected at Station M1 of the Monterey Bay Time-series. A.) Contour plot of temperature over the study period. Black dots represent profiling CTD data point locations used for extrapolation. B.) QPCR abundances of *Bathycoccus* clade BI and *Ostreococcus* clade OI at the surface (<5 m) and deeper in the water column (10 m or 20 m) representing subsurface chlorophyll max, if present, or the base of the wind-mixed surface layer. C.) *Bathycoccus* clade BII and *Ostreococcus* clade OII detected over the study period in the same samples as for B.). Note the different Y-axis scales in B.) and C.). n.d., clade not detected by qPCR.

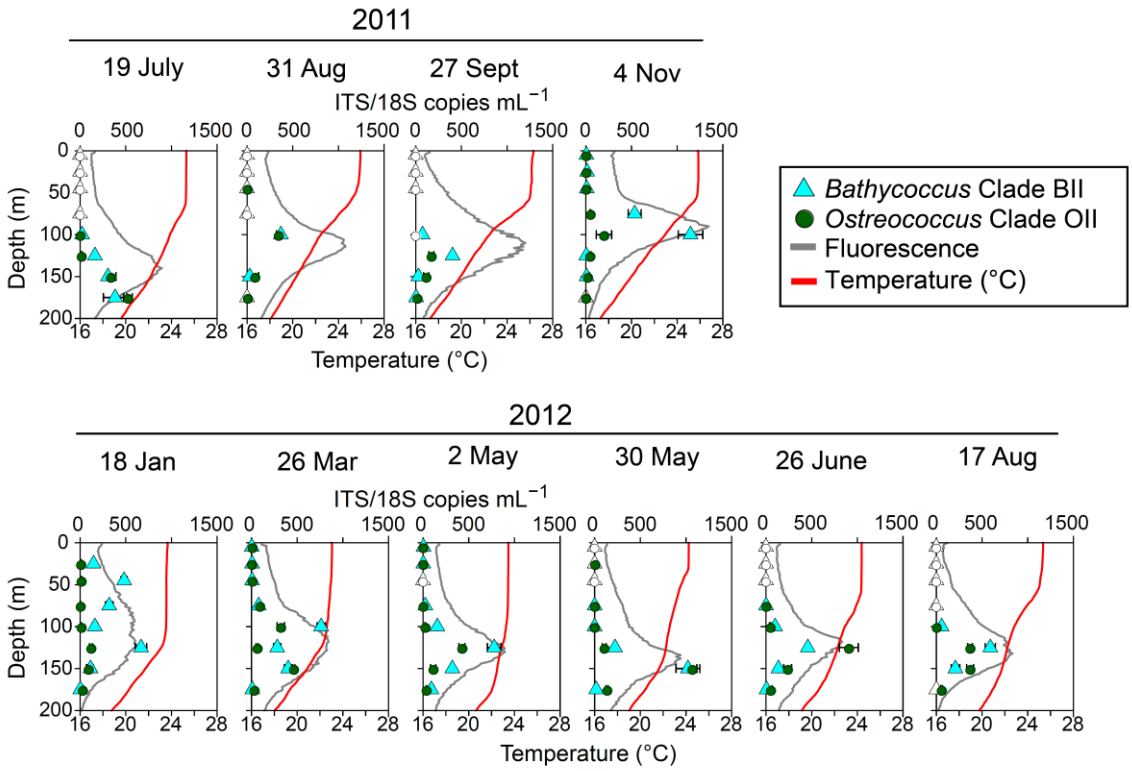


Figure 5. Profiles of temperature, fluorescence, and qPCR abundance sampled over a 14-month period at Station ALOHA of the Hawaii Ocean Time-series. Only the OII (green) and BII (blue) clades were detected by qPCR. Open symbols represent samples where the clade was not detected. Error bars represent standard deviation of triplicate qPCR reactions.

	BII present	BII absent	OII present	OII absent
Depth (m)	10 ± 11	6 ± 7	10 ± 11	6 ± 7
Temperature (°C)	15.7 ± 1.0	13.4 ± 2.0	15.8 ± 1.1	13.4 ± 1.9
Salinity	33.31 ± 0.11	33.49 ± 0.21	33.36 ± 0.15	33.47 ± 0.21
NH4+	0.06 ± 0.08	0.25 ± 0.25	0.07 ± 0.07	0.25 ± 0.25
NO_3^-	1.22 ± 2.29	4.57 ± 3.26	1.20 ± 2.31	4.58 ± 3.24
PO_4^{3-}	0.59 ± 0.31	0.86 ± 0.34	0.58 ± 0.32	0.86 ± 0.33
n=	7	15	7	15

Table 5. Mean and standard deviation of environmental variables at Station M1 (MBTS) in samples where *Bathycoccus* BII or *Ostreococcus* OII were detected, and where BII and OII were absent.

	BII & OII present	BII & OII absent
Depth (m)	107.7 ± 51.8	37.5 ± 28.4
Temperature (°C)	22.4 ± 1.6	24.8 ± 1.1
NO_3^-	0.36 ± 0.54	0.02 ± 0.01
PO_4^{3-}	0.10 ± 0.04	0.09 ± 0.04
n=	49	22

Table 6. Mean and standard deviation of environmental variables at Station ALOHA in samples where both *Bathycoccus* BII and *Ostreococcus* OII were detected, and where neither genus was observed.

Biogeography of Bathycoccus and Ostreococcus Clades

To broaden results from the temporal studies at Station M1 and Station ALOHA we examined samples from other locations (Fig. 3, Table 1, Supplementary Datafile). These included a transect cruise (WFAD09) along the California Cooperative Oceanic Fisheries Investigations Line 67 from coastal Monterey Bay southwestward, crossing the core of the California Current and ending approximately 800 km offshore

at the edge of oligotrophic North Pacific Subtropical Gyre waters (Table 1, Supplementary DataFile). In addition a cruise was performed that originated in coastal Monterey Bay and went southward along the coast of North America, terminating in tropical waters approximately 200 km south of Baja California, Mexico (GOC12, Fig. 6). And samples from other cruises along Line 67, Line P, in the Kuroshio Current and the tropical Atlantic were also investigated.

During cruise WFAD09, both *Bathycoccus* clades were detected at all stations sampled along Line 67, except clade BII was not detected at the Monterey Bay station closest to shore (~5 km, Fig. 7). The abundance of clade BI decreased moving westward as waters became increasingly oligotrophic. Clade BII abundances were comparable between a transition zone (between stations 67-65 and 67-95) and oligotrophic stations (stations 67-145 and 67-155). Distributions of *Ostreococcus* clades OI and OII from this cruise were reported previously (Simmons et al 2016), and showed more separation in terms of sampling location than *Bathycoccus*. *Ostreococcus* clade OI was present at a maximum of 14,027 ± 548 18S copies mL^{-1} near the coast and in the transition zone, but was not detected farther offshore (Fig. 7). *Ostreococcus* clade OII was observed in the DCM at the two most oligotrophic stations (maximum abundance 72 ± 3 18S copies mL^{-1}, 100 m depth), but otherwise only at the westward edge of the transition zone (station 67-95).

The GOC12 transect also began in the cool, upwelling-influenced coastal environment of Monterey Bay where NO_3^- and PO_4^{3-} were 4.48 and 0.61 µM, respectively, at the surface (<5 m) of the coldest station sampled (12.0°C;

Supplemental Datafile). As the cruise progressed southwards, it transitioned into warmer waters that reached 24.2°C at the southernmost station, south of Baja California, Mexico, where surface NO_3^- was below the 0.01 µM detection limit and PO_4^{3-} was 0.28 µM (Supplemental Datafile). Both genera were detected at every station sampled, however *Bathycoccus* clade BI and *Ostreococcus* clade OI were not detected at the southernmost station, and *Bathycoccus* clade BII and *Ostreococcus* clade OII were not detected at the two northernmost stations (Fig. 6). In between the transect end points was a large region where all four clades co-existed. Because methodologies (including for metadata) were identical for the WFAD09 and GOC12 cruises performed on the *R/V Western Flyer* we explored characteristics where clades of each genus co-occurred in the same samples. Ratios of *Bathycoccus* clade BI to *Bathycoccus* clade BII (clade BI/clade BII) were examined in the 53 WFAD09 and GOC12 samples where both were detected (out of 74 total; Fig. 8). It should be noted that water column structure did vary in these samples, with the more nutrient rich coastal samples typically having shallower mixed layers than those further south (GOC12) or west (WFAD09). Depth ($\rho = -0.74$), PO_4^{3-} ($\rho = -0.48$) and salinity ($\rho = -0.43$) were significantly correlated (Spearman's rank correlation coefficient) to the *Bathycoccus* clade ratios. This analysis was also performed with the 22 samples where the *Ostreococcus* clades co-occurred (clade OI/clade OII; Fig. 8), revealing that depth ($\rho = -0.69$), PO_4^{3-} ($\rho = -0.64$), salinity ($\rho = -0.59$) and NO_3^- ($\rho = -0.44$) were significantly correlated to the clade ratios.

Clade abundances also varied in the other samples analyzed. *Bathycoccus* BII and *Ostreococcus* OII were not detected in Line P samples (Table 1). Moving from coastal to offshore, subpolar gyre waters along Line P, *Bathycoccus* BI and *Ostreococcus* OI abundances increased from 803 and 625 ITS/18S copies mL^{-1} (Station P1) to 10,667 and 35,797 ITS/18S copies mL^{-1} (Station P6), respectively, and BI declined again at the most westward station (Station P8, 1,405 ± 73 ITS copies mL^{-1}) where OI abundances were an order of magnitude higher (13,766 ± 1,843 18S copies mL^{-1}) than BI. Cruises CANON11, S410, and MV1405 sampled the California Current System (Table 1) and were dominated by *Bathycoccus* clade BI and *Ostreococcus* clade OI, while clades BII and OII were only detected at the southernmost of the three stations sampled during MV1405, which occurred during the 2014 period when the anomalously warm patch was present in the region. In the frontal zone between the Kuroshio and Oyashio Currents, high abundances of *Ostreococcus* clade OII (50,189 ± 561 18S copies mL^{-1}) were reported, with OI also present but only to a maximum of 6,825 ± 198 18S copies mL^{-1} (Clayton et al 2016). For *Bathycoccus* in these same samples, we observed maximum abundances of 1,751 ± 294 ITS copies mL^{-1} for BI and 2,662 ± 318 ITS copies mL^{-1} for BII. Lastly, *Ostreococcus* OII was previously reported to be below detection at the majority of tropical Atlantic samples analyzed, with a maximum of 73 18S copies mL^{-1} (Demir-Hilton et al 2011). In contrast, we observed *Bathycoccus* BII in all tropical Atlantic samples analyzed and it reached a maximum of 4,125 ITS copies mL^{-1} at a depth of 61 m where NO_3^- was 0.03 µM and PO_4^{3-} was not detected (<0.01 µM). Neither OI

nor BI were detected in this large transect and OII was present in one sample (Supplemental Datafile). Taken together, these samples illustrated a general trend of OI and BI being present in cooler (7.8-24.1^{o}C and 7.8-25.2 ^{o}C, respectively), lower salinity waters (30.45-35.06), and OII and BII being present in warmer (11.3-27.3^{o}C and 9.9-28.9^{o}C, respectively), higher salinity waters (33.03-36.39 and 32.72-36.39, respectively). Additionally, both *Ostreococcus* clade OII and *Bathycoccus* clade BII were found over a wide range of depths, from the surface to 175 m, with the latter being the deepest depth sampled in this study.

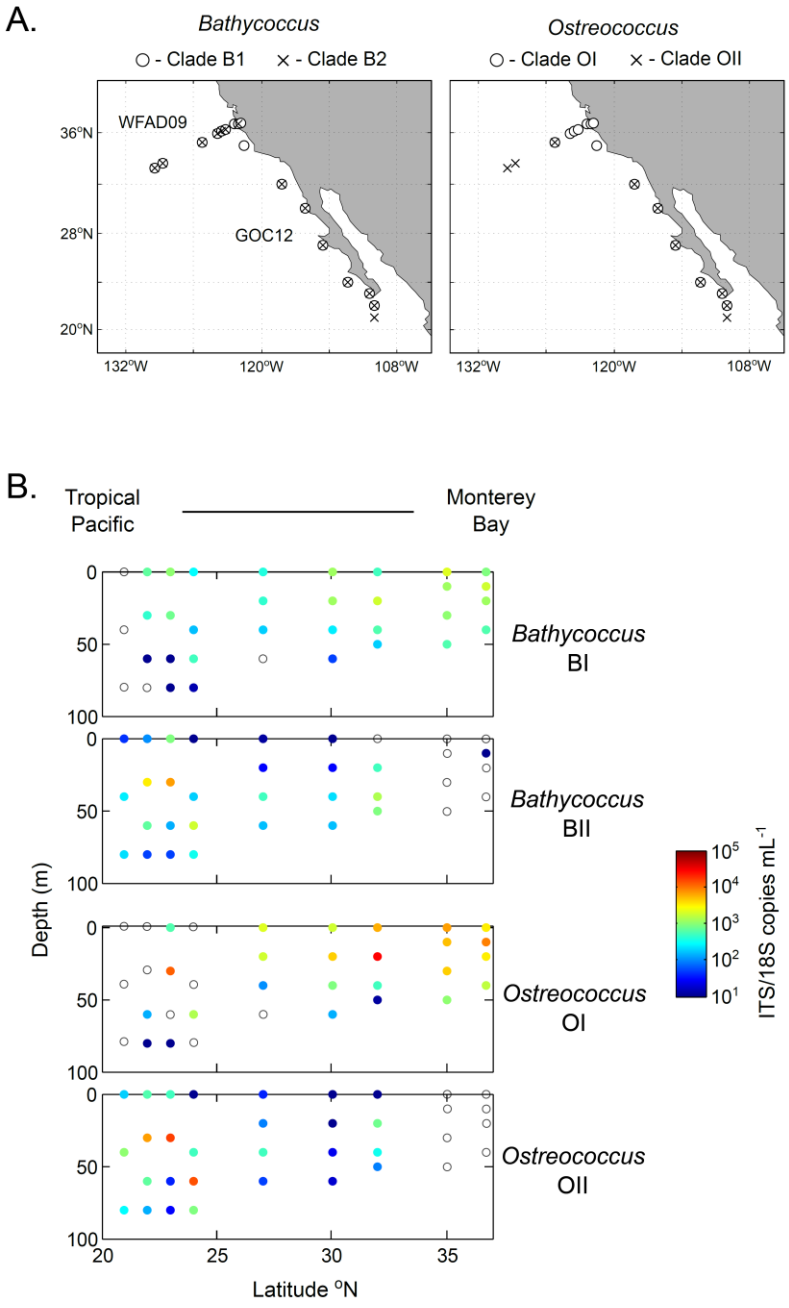


Figure 6. Abundances of *Bathycoccus* and *Ostreococcus* clades on the GOC12 cruise transect (February 2012). A.) Locations of stations where OI or BI clades were detected (open circle) and where the OII or BII clades were detected (x) are indicated. B.) Abundances of *Bathycoccus* and *Ostreococcus* clades by qPCR at different depths at each station during the transect cruise. Note the logarithmic scaling for abundances (heat map).

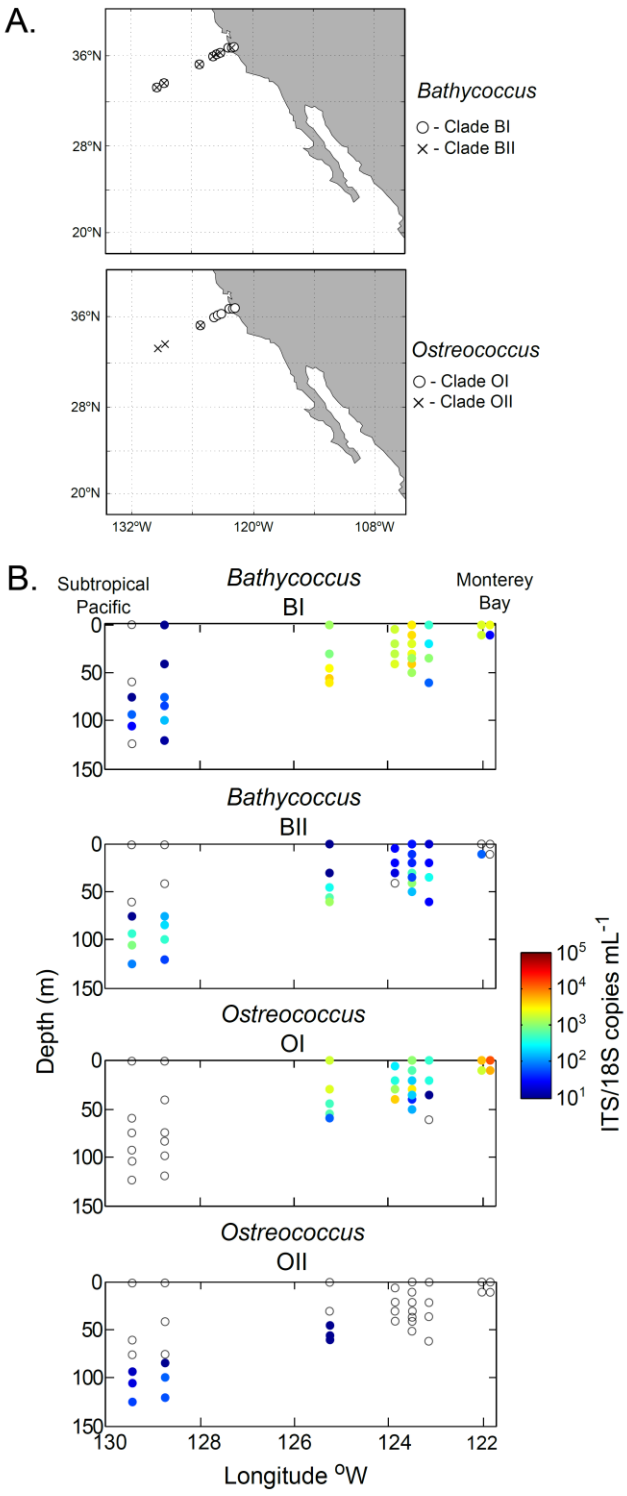

A.
36°N
28°N
20°N
Bathycoccus
○ - Clade BI
× - Clade BII
36°N
28°N
20°N
Ostreococcus
○ - Clade OI
× - Clade OII
132°W
120°W
108°W
B.
Subtropical Pacific
Bathycoccus BI
Monterey Bay
Bathycoccus BII
Ostreococcus OI
Ostreococcus OII
Depth (m)
0
50
100
150
130
128
126
124
122
Longitude °W
10^5
10^4
10^3
10^2
10^1
ITS/18S copies mL^-1

Figure 7. Abundances of the *Bathycoccus* and *Ostreococcus* clades on the WFAD09 transect (Line 67) in the eastern North Pacific. A.) *Ostreococcus* and/or *Bathycoccus* were present at all stations samples during October 2009. Stations where OI or BI clades were detected (open circle) and where the OII or BII clades were detected (x) are indicated. B.) Profiles of *Bathycoccus* and *Ostreococcus* clade abundances as determined by qPCR. Note the logarithmic scaling for abundances (heat map).

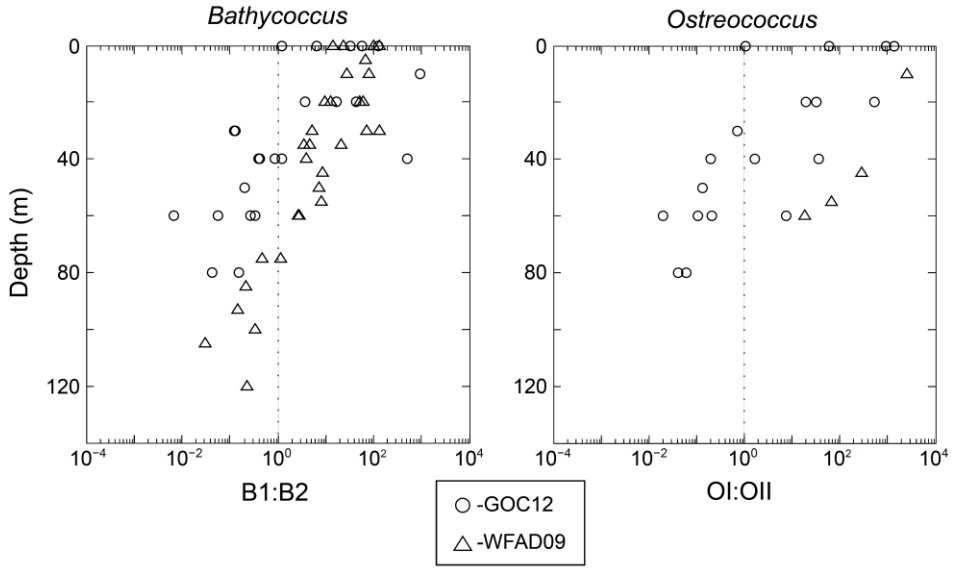


Figure 8. Clade ratios (*Bathycoccus* BI:BII or *Ostreococcus* OI:OII) of abundances determined by qPCR at stations where ecotypes co-occurred on WFAD09 and GOC12 cruises. Ratios are plotted against depth on y-axis. Note that ecotype ratios are shown on a logarithmic scale.

Statistical Analyses Using Data from Across Multiple Regions

We performed additional tests on the qPCR and environmental data to search for generalizable trends in a dataset comprised of the 230 samples where at least one clade of *Bathycoccus* or *Ostreococcus* was detected. Since qPCR and environmental data failed tests for normality, non-parametric statistical tests were chosen to assess

significant relationships. The complete sample set was also heavily skewed towards surface and coastal samples and given the breadth of water columns sampled, associations with depth also reflect differences in water column structure, irradiance, nutrient availability and other factors. Moreover, the various nitrogen sources measured (NO_3^-, NO_2^- or NH_4^+) were often below detection while PO_4^{3-} was detectable in all Pacific Ocean samples. Clades BI and OI showed positive correlations with PO_4^{3-} concentrations ($\rho = 0.58$ and $\rho = 0.51$, respectively), while clades BII and OII were negatively correlated to inorganic nutrients (for PO_4^{3-}, $\rho = -0.30$ and $\rho = -0.44$, respectively), due to their more frequent presence in stratified habitats. A significant positive correlation was observed for *Bathycoccus* clade BII with depth (median depth of 40 m, Spearman's correlation $\rho = 0.25$) and for *Ostreococcus* clade OI with chlorophyll *a* concentrations ($\rho = 0.49$). Temperature and salinity were negatively correlated with abundance of *Bathycoccus* clade BI ($\rho = -0.65$ and $\rho = -0.76$) and *Ostreococcus* clade OI ($\rho = -0.63$ and $\rho = -0.65$), reflecting their presence in the relatively cool, lower salinity waters of the Northeast Pacific, California Current System and upwelling-influenced coastal environments (Fig. 9, Table 7). In contrast, temperature and salinity were positively associated with *Bathycoccus* clade BII ($\rho = 0.43$ and $\rho = 0.42$) and *Ostreococcus* clade OII ($\rho = 0.57$ and $\rho = 0.61$) which were found in tropical, subtropical, and warmer temperate environments (Fig. 9, Table 7).

To further examine how *Bathycoccus* and *Ostreococcus* clade abundances varied with environmental conditions, a multivariate transformation of physico-chemical

parameters was combined with a clustering analysis to define habitat types. Since the sample set had uneven representation of different sites and types of environments, non-metric multidimensional scaling (NMDS) analysis was used to transform a dissimilarity matrix containing temperature, salinity, depth, PO_4^{3-}, NO_3^-, and N:P data for samples where at least one clade was detected by qPCR. K-means clustering partitioned samples into three groups (Fig. 10a) which were then used to bin qPCR abundance data and further describe the biogeography of each clade (Fig. 10b, 10c). Cluster 1 contained samples from the North Pacific Subtropical Gyre and tropical Atlantic (latitudinal range: 2° to 36° N), typically characterized by warm (23.1 ± 2.0°C), nutrient-poor (0.51 ± 0.67 µM and 0.11 ± 0.09 µM NO_3^- and PO_4^{3-}, respectively), stratified, deep photic zone environments like ALOHA (median depth 76 m), but most Kuroshio samples grouped here as well. Not only was Cluster 1 dominated by clades BII and OII, but here they had their abundance maximum relative to the other two clusters (Fig. 10). In contrast, BI and OI were not detected in Cluster 1 samples, except for in some Kuroshio samples. The other two clusters partitioned temperate region samples (latitudinal range: 21° to 49°N) into lower- (Cluster 2) and higher- (Cluster 3) nutrient samples, and contained higher abundances of clades BI and OI than Cluster 1 (Fig. 10). The temperate, lower-nutrient group (Cluster 2) contained samples with temperatures of 14.9 ± 3.0°C from shallower photic zones wherein the median depth sampled was 10 m and N:P ratio was 2.6 ± 3.0. Cluster 3, the temperate, higher-nutrient group (12.2 ± 2.6°C, median depth 40 m, 7.7 ± 2.2 [N:P] for Cluster 3, $p<0.01$), had higher median abundances of all four

ecotypes than Cluster 2 (Fig. 10). Abundance of *Bathycoccus* and *Ostreococcus* clades across the three environmental clusters were also compared by Kruskal-Wallace tests with Dunn's post-hoc tests. No clade showed a significant difference in abundance between Clusters 2 and 3 using this statistical test ($p>0.01$). However, the abundance of clades BI and OI were significantly ($p<0.01$) lower in Cluster 1 than in Clusters 2 and 3. Moreover, clades BII and OII were significantly higher in Cluster 1 compared to the other groups.

	Bathycoccus		***Ostreococcus***	
	BI	**BII**	**OI**	**OII**
Temperature (°C)	-0.65	0.43	-0.63	0.57
Salinity	-0.76	0.42	-0.65	0.61
Depth (m)	-0.43	0.25	-0.45	0.16*
NO_3^- (µM)	0.35	-0.18	0.39	-0.17*
PO_4^{3-} (µM)	0.58	-0.30	0.51	-0.44
Chlorophyll a	0.27*	-0.27*	0.49	-0.28*

Table 7. Spearman's rank correlation coefficient (ρ) between ecotype qPCR abundances and environmental variables. * indicates correlation not significant (ρ >0.01).

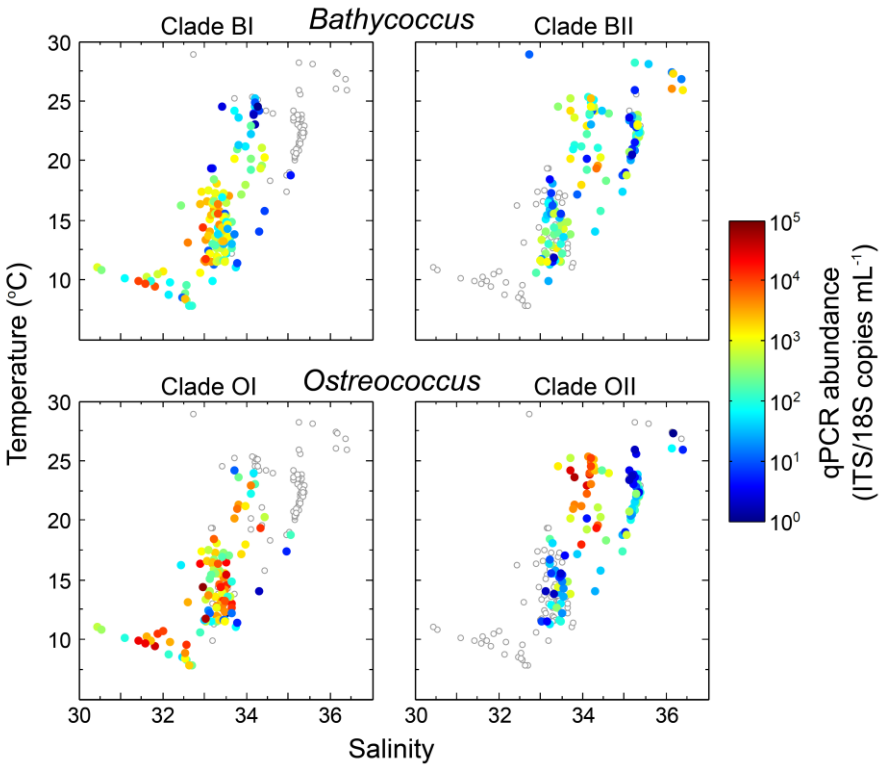
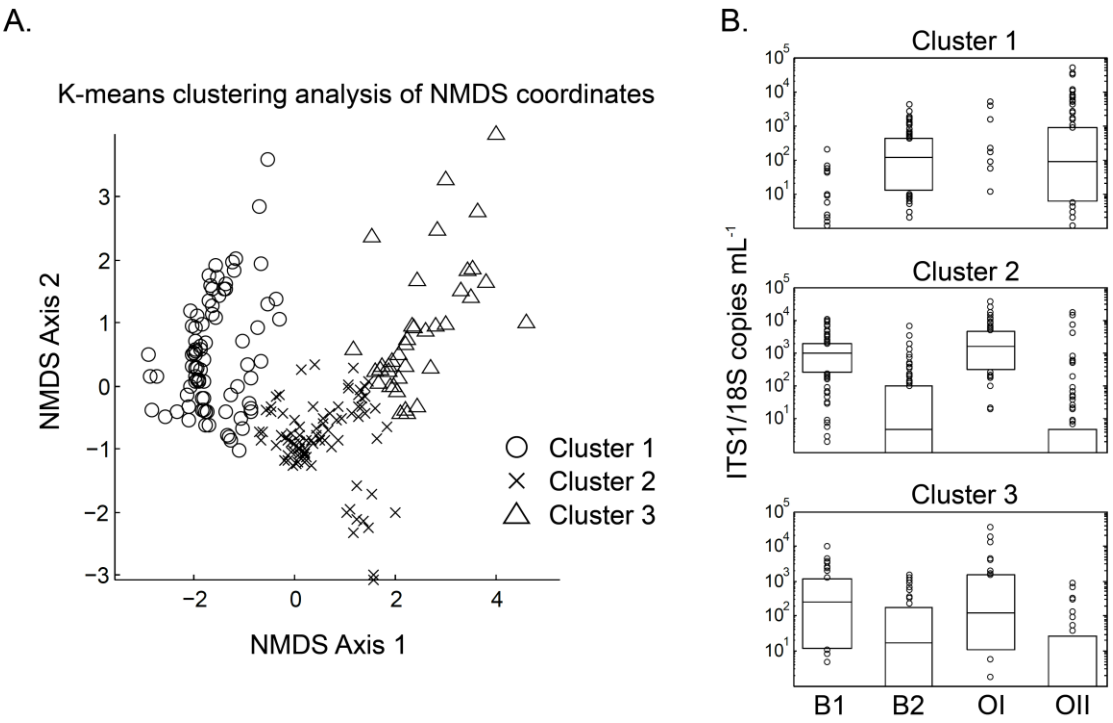


Figure 9. Abundance of *Bathycoccus* and *Ostreococcus* clades determined by qPCR in relation to temperature and salinity. *Bathycoccus* clades BI and BII were detected in 152 and 153 samples, respectively, and *Ostreococcus* clades OI and OII in 131 and 124 samples, respectively. Thus, at least one of these taxa was present in 230 of the 266 total samples.



C.

	Depth (m)	Temp (°C)	Salinity (ppt)	NO_3^- (µM)	PO_4^{3-}(µM)	N:P
Cluster 1	74 ± 65	23.1 ± 2.0	34.98 ± 0.59	0.51 ± 0.67	0.11 ± 0.09	5.6 ± 8.9
Cluster 2	20 ± 26	14.9 ± 3.0	33.15 ± 0.73	1.49 ± 1.84	0.54 ± 0.23	2.6 ± 3.0
Cluster 3	43 ± 26	12.2 ± 2.6	33.41 ± 0.63	10.46 ± 4.49	1.35 ± 0.37	7.7 ± 2.2

Figure 10. A.) Nonmetric multidimensional scaling coordinates of physical and chemical parameters were used for k-means clustering analysis. B.) Boxplots showing abundances of each clade by cluster. The box represents the first quartile, median, and third quartile. Values below first quantile and above third quantile are also shown (circles). C.) Table of mean and standard deviation for various environmental variables associated with each clade.

Discussion

Prior phylogenetic analysis of 5.8S/ITS2 environmental data from Mamiellophyceae delineated two clades within the genus *Bathycoccus*, which have identical 18S rRNA gene sequences (Monier et al 2013). One of these was represented in culture and field samples (clade BI) while the other (clade BII) was only found in environmental samples. ITS sequences are less conserved than rRNA genes such as 18S, thus the findings suggested that *Bathycoccus* clades BI and BII are more closely related than *Ostreococcus* clades OI (represented by genome sequenced isolate *Ostreococcus lucimarinus*) and OII (represented by genome sequenced isolate RCC809), which exhibited divergence in both the 18S rRNA gene and the ITS (Marin and Melkonian 2010, Rodriguez et al 2005). Comparison of the BII clade, represented by a targeted metagenome assembly where multiple cells were sorted and sequence from a single tropical Atlantic population, and the cultured BI clade (strain RCC1105) genome (Moreau et al. 2012), as well as genomes from cultures of OI and OII, and field metatranscriptomes, have underscored the finding that the two *Bathycoccus* clades are less diverged than OI and OII. Nevertheless, the extent of BI and BII divergence is notable and led to the proposal that the two *Bathycoccus* ITS clades represent distinct ecotypes (Monier et al 2013, Simmons et al 2016), akin to the mesotrophic (OI) and oligotrophic (OII) ecotypes of *Ostreococcus* established previously (Demir-Hilton et al 2011). Mapping of bulk metagenomic data from TARA Oceans against a newer targeted metagenome (where single cells were

presumed to be from the same population and combined) of BII and the RCC1105 genome also led to the conclusion that two ecotypes exist (Vannier et al. in press). Specifically, the latter study concluded that the two clades "…rarely co-occur and occupy distinct oceanic niches in particular with respect to depth."

With respect to quantitative studies, *Ostreococcus* and *Bathycoccus* cells have been enumerated by FISH or qPCR in mesotrophic, coastal ecosystems (Collado-Fabbri et al 2011, Countway and Caron 2006, Not et al 2004, Not et al 2005) and oligotrophic, oceanic environments (Gomez-Pereira et al 2013, Marie et al 2006, Not et al 2008, Treusch et al 2012, Zhu et al 2005). However, fewer studies have been performed that discriminate the two *Ostreococcus* ecotypes (Clayton et al 2016, Demir-Hilton et al 2011, Simmons et al 2016) and to date none quantified the two *Bathycoccus* clades as distinct entities.

Here, analyses of ITS1/5.8S/ITS2 environmental clone libraries exhibited two *Bathycoccus* ITS clades, as seen in previous studies. Given a preliminary sense of distributions from prior studies, we searched culture collections for isolates from locations where the BII clade seemed to thrive and were able to identify cultured representatives of clade BII (RCC715 and RCC716; Fig. 3B) isolated from the Indian Ocean at a depth of 70 m (http://roscoff-culture-collection.org/rcc-strain-details/716). With a greater number of *Bathycoccus* ITS sequences in hand we designed two effective qPCR assays for discriminating and quantifying ITS1 sequences from clades BI and BII. The two new *Bathycoccus* clade specific qPCR assays and pre-existing assays for *Ostreococcus* ecotypes OI and OII resolved distributions and abundances

of these clades on a set of 266 photic zone samples, 230 of which contained *Bathycoccus*, *Ostreococcus* or both. *Bathycoccus* clade BI and *Ostreococcus* clade OI were associated with cooler temperatures, lower salinity, higher nutrient conditions that typically occur in more coastal (shallower) photic zones. In contrast, *Bathycoccus* clade BII and *Ostreococcus* clade OII were found in environments with warmer temperatures, higher salinity, lower nutrients, deeper photic zones, and lower chlorophyll concentrations (Figs. 9, 10, Table 7).

Prior field research led to the proposal that *Ostreococcus* clades OI and OII represent ecotypes adapted to mesotrophic and oligotrophic conditions, respectively (Demir-Hilton et al 2011). Given the extensive geographic overlap and co-occurrence of *Bathycoccus* clade BI and *Ostreococcus* clade OI, we propose that they indeed both represent ecotypes adapted to mesotrophic environmental conditions. Similarly, overlapping distributions were observed between *Bathycoccus* clade BII and *Ostreococcus* clade OII, therefore we propose that they represent oligotrophic-adapted ecotypes. However, important differences in biogeography were observed among the clades of each genera. *Bathycoccus* clades BI and BII were observed at a wider range of environmental conditions and overlapped in a greater number of samples (85 samples) than clades OI and OII did (42 samples). This suggests that while *Ostreococcus* OII is adapted to oligotrophic conditions found in oceanic settings (apart from highly oligotrophic surface waters in strongly stratified water columns), that BII tolerates a wider range of environmental conditions. The higher rate of co-existence between *Bathycoccus* clades could also reflect the observation

that *Bathycoccus* clades appear to be less evolutionarily divergent than *Ostreococcus* clades (Simmons et al 2016), therefore the ecological niches of *Bathycoccus* clades might overlap to a greater extent as well.

Regions where *Bathycoccus* and *Ostreococcus* ecotypes co-occur also revealed important factors with respect to parameters that might drive distributions of these algae. Previous investigations have suggested that *Ostreococcus* OI is a 'high-light' adapted clade, whereas *Ostreococcus* OII is 'low-light' adapted, similar to ecotypes of *Prochlorococcus* (Johnson et al 2006). Clades BII and OII were found in relatively high abundances over a wide range of depths, from the surface (e.g., Kuroshio Front, and less stratified periods at ALOHA) to depths of 175 m. Thus, these clades represent 'light-polyvalent' ecotypes as opposed to 'low-light' adapted taxa. Using data from the two cruises with the best correspondence in terms of sampling strategy and methodologies (WFAD09 and GOC12, Table 1), we could not discriminate between depth, phosphate or nitrate as potential drivers for greater relative abundances of *Bathycoccus* BI and *Ostreococcus* OI at shallower depths versus clades *Bathycoccus* BII and *Ostreococcus* OII which represented a larger fraction of the community deeper in the water column (Fig. 8). The Spearman's correlation coefficient was notably stronger for PO_4^{3-}, NO_3^-, and salinity for *Ostreococcus* OI/OII ratios, and weaker for depth, than seen for the *Bathycoccus* BI/BII ratios which exhibited significant relationships for depth, PO_4^{3-} and salinity. This highlights a stronger delineation in the salinity and phosphate characteristics of waters

containing *Ostreococcus* OII versus *Ostreococcus* OI (and weaker connection to depth) as compared to those containing *Bathycoccus* BI or BII.

The two time-series studies' considered here revealed temporal dynamics of the *Bathycoccus* and *Ostreococcus* clades. In the MBTS dataset, our study period serendipitously captured an unusual oceanographic phenomenon when a strong warm anomaly entered Monterey Bay, bringing with it record high sea surface temperatures. This anomalously warm patch of water appeared in the Northeast Pacific at the end of 2013, and moved eastward towards the North American coast May-September 2014, and persisted into 2015 (Bond et al 2015, Kintisch 2015). Satellite-based remote-sensing revealed that this warm patch caused the movement of the high productivity frontal region between the subarctic and subtropical North Pacific ~300 km northward during winter 2013/2014 (Whitney 2015). This resulted in an expansion of warm, low productivity subtropical conditions characterized by low surface concentrations of chlorophyll *a* in much of the north eastern Pacific Ocean alongside a deepening of the nutricline and subsurface chlorophyll maximum in the California Current System (Zaba and Rudnick 2016). These changes have been speculated to have resulted in reduced phytoplankton productivity and shifts in the phytoplankton community (Zaba and Rudnick 2016). In our MBTS dataset, the appearance of a strong warm anomaly on the 12 August sampling date was accompanied by a decrease in nutrient stocks and a shift from *Ostreococcus* OI to *Bathycoccus* BI as the most abundant taxa (of those enumerated herein). In addition, it led to the first report of *Bathycoccus* BII and *Ostreococcus* OII (this study) in this

location, while several prior studies did not observed these ecotypes in Monterey Bay or nearby (Demir-Hilton et al 2011, Ottesen et al 2013, Simmons et al 2016). Prior to the arrival of the warm anomaly in Monterey Bay during the spring/summer upwelling period, abundances of *Bathycoccus* BI and *Ostreococcus* OI were low on 5 May when the influence of wind-driven coastal upwelling was strongest, as indicated by cooler sea surface temperatures and high nutrient and chlorophyll a concentrations. Low abundances of picoeukaryotes during periods of upwelling have also been reported in previous studies off the coasts of Oregon and Chile (Collado-Fabbri et al 2011, Sherr et al 2005). As the nutrient-rich water column began to stratify in June, *Ostreococcus* OI experienced a "bloom" (18,033 ± 1,260 18S copies mL^{-1}) while *Bathycoccus* BI abundances stayed low (31 ± 6 ITS copies mL^{-1}) (Fig. 4). *Bathycoccus* BII reached maximum abundance (8,444 ± 529 ITS copies mL^{-1}) on 20 November, at which point *Ostreococcus* OI abundance was relatively low (1,877 ± 69 18S copies mL^{-1}), indicating that these taxa respond to different environmental conditions. In addition, clades BII and OII arrived with the warm anomaly. In our broader dataset, these clades are associated with warmer, low-nutrient conditions, suggesting that the warm anomaly underlay the observed changes in Mamiellophyceae phytoplankton community structure. It remains unclear whether these community shifts were caused by changes in temperature, nutrient stocks, or other associated factors.

In contrast to MBTS, our time-series at Station ALOHA displayed less environmental variability, characteristic of this study location in the oligotrophic

NPSG. *Bathycoccus* BII and *Ostreococcus* OII persisted year-round in the DCM, and were observed in the surface layer during periods of enhanced wind-mixing. Maximum abundances of these taxa occurred during and after the winter wind-mixing period, in agreement with observations from the Sargasso Sea at BATS where the timing of Mamiellophyceae blooms was correlated to deep winter mixing events (Treusch et al 2012). However, during strongly stratified periods in the Sargasso Sea, *Ostreococcus* disappeared from the water column and only *Bathycoccus* remained above the qPCR detection limits (Treusch et al 2012). By comparison, at Station ALOHA where phosphate is never as deplete as at BATS during the stratified summer period, BII and OII were observed at the DCM in relatively high abundances throughout the stratified summer months (Fig. 5). Since oligotrophic subtropical gyres cover vast regions of the planet and represent a significant carbon sink, the general persistence of *Ostreococcus*, and especially *Bathycoccus*, in gyre habitats indicates that they are important contributors to primary production throughout the year.

A study by Dutkiewicz et al (2009) explored the concept of r and K strategists in phytoplankton with an *in-silico* ecosystem model that utilized virtual phytoplankton types with stochastically assigned traits to create "self-assembling" communities based on competition for resources, ocean circulation models and biogeochemical cycling in different regions of the globe. Dutkiewicz and colleagues found that higher latitude oceans experience a large degree of environmental disturbance due to significant seasonal variability, which favors the r strategists due to their ability to

grow rapidly when environmental conditions are optimal and nutrient concentrations are high (i.e. “bloomers”). Conversely, the tropical and subtropical seas are characterized by weak seasonality and relatively stable conditions, and are dominated by K strategists due to their relatively low nutrient quotas and slow growth rates (i.e. “gleaners”). In the context of *Bathycoccus* and *Ostreococcus* ecotypes, clades BI and OI would represent r strategists due to their presence in high latitude seas that experience seasonal mixing and have comparatively high primary productivity. *Ostreococcus* clade OII and *Bathycoccus* clade BII would constitute K strategists since they inhabit low latitude, nutrient-poor, stratified habitats that experience little seasonal variability, although *Bathycoccus* clade BII falls less clearly into either category.

References

Bond NA, Cronin MF, Freeland H, Mantua N (2015). Causes and impacts of the 2014 warm anomaly in the NE Pacific. *Geophys Res Let* **42:** 3414-3420.

Bustin SA (2004). *AZ of quantitative PCR*. International University Line: La Jolla, CA USA.

Clayton S, Lin YC, Follows MJ, Worden AZ (2016). Co-existence of distinct Ostreococcus ecotypes at an oceanic front. *Limnol Oceanogr* doi: 10.1002/lno.10373

Coleman AW (2009). Is there a molecular key to the level of "biological species" in eukaryotes? A DNA guide. *Mol Phylogenet Evol* **50:** 197-203.

Collado-Fabbri S, Vaulot D, Ulloa O (2011). Structure and seasonal dynamics of the eukaryotic picophytoplankton community in a wind-driven coastal upwelling ecosystem. *Limnol Oceanogr* **56:** 2334-2346.

Countway PD, Caron DA (2006). Abundance and distribution of Ostreococcus sp. in the San Pedro Channel, California, as revealed by quantitative PCR. *Appl Environ Microbiol* **72:** 2496-2506.

de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R *et al.* (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science* **348:** 1261605.

Demir-Hilton E, Sudek S, Cuvelier ML, Gentemann C, Zehr JP, Worden AZ (2011). Global distribution patterns of distinct clades of the photosynthetic picoeukaryote *Ostreococcus*. *ISME J* **5:** 1095-1107.

Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S *et al.* (2006). Genome analysis of the smallest free-living eukaryote Ostreococcus tauri unveils many unique features. *Proc Natl Acad Sci U S A* **103:** 11647-11652.

Dutkiewicz S, Follows MJ, Bragg JG (2009). Modeling the coupling of ocean ecology and biogeochemistry. *Global Biogeochem Cycles* **23**(4).

Gomez-Pereira PR, Kennaway G, Fuchs BM, Tarran GA, Zubkov MV (2013). Flow cytometric identification of Mamiellales clade II in the Southern Atlantic Ocean. *FEMS Microbiol Ecol* **83:** 664-671.

Guillard R, Hargraves P (1993). Stichochrysis immobilis is a diatom, not a chrysophyte. *Phycologia* **32:** 234-236.

Guillou L, Eikrem W, Chretiennot-Dinet M, Le Gall F, Massana R, Romari K *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:** 193-214.

Holmes RM, Aminot A, Kérouel R, Hooker BA, Peterson BJ (1999). A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Can J Fish Aquat Sci* **56:** 1801-1808.

Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EMS, Chisholm SW (2006). Niche partitioning among Prochlorococcus ecotypes along ocean-scale environmental gradients. *Science* **311:** 1737-1740.

Katoh K, Standley DM (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol biol and evol* **30:** 772-780.

Kintisch E (2015). 'The Blob'invades Pacific, flummoxing climate experts. *Science* **348:** 17-18.

Marie D, Zhu F, Balague V, Ras J, Vaulot D (2006). Eukaryotic picoplankton communities of the Mediterranean Sea in summer assessed by molecular approaches (DGGE, TTGE, QPCR). *FEMS Microbiol Ecol* **55:** 403-415.

Marin B, Klingberg M, Melkonian M (1998). Phylogenetic relationships among the Cryptophyta: analyses of nuclear-encoded SSU rRNA sequences support the monophyly of extant plastid-containing lineages. *Protist* **149:** 265-276.

Marin B, Melkonian M (2010). Molecular phylogeny and classification of the Mamiellophyceae class. nov. (Chlorophyta) based on sequence comparisons of the nuclear- and plastid-encoded rRNA operons. *Protist* **161:** 304-336.

Moisander PH, Beinart RA, Voss M, Zehr JP (2008). Diversity and abundance of diazotrophic microorganisms in the South China Sea during intermonsoon. *ISME J* **2:** 954-967.

Monier A, Welsh RM, Gentemann C, Weinstock G, Sodergren E, Armbrust EV *et al.* (2012). Phosphate transporters in marine phytoplankton and their viruses: cross-domain commonalities in viral-host gene exchanges. *Environ Microbiol* **14:** 162-176.

Monier A, Sudek S, Fast NM, Worden AZ (2013). Gene invasion in distant eukaryotic lineages: discovery of mutually exclusive genetic elements reveals marine biodiversity. *ISME J* **7:** 1764-1774.

Monier A, Worden AZ, Richards TA (2016). Phylogenetic diversity and biogeography of the Mamiellophyceae lineage of eukaryotic phytoplankton across the oceans. *Environ Microbiol Rep*. **8**(4): 461-469.

Moore C, Mills M, Arrigo K, Berman-Frank I, Bopp L, Boyd P *et al.* (2013). Processes and patterns of oceanic nutrient limitation. *Nat Geosci* **6:** 701-710.

Moreau H, Verhelst B, Couloux A, Derelle E, Rombauts S, Grimsley N *et al.* (2012). Gene functionalities and genome structure in Bathycoccus prasinos reflect cellular specializations at the base of the green lineage. *Genome Biol* **13:** R74.

Not F, Latasa M, Marie D, Cariou T, Vaulot D, Simon N (2004). A single species, Micromonas pusilla (Prasinophyceae), dominates the eukaryotic picoplankton in the Western English Channel. *Appl Environ Microbiol* **70:** 4064-4072.

Not F, Massana R, Latasa M, Marie D, Colson C, Eikrem W *et al.* (2005). Late summer community composition and abundance of photosynthetic picoeukaryotes in Norwegian and Barents Seas. *Limnol Oceanogr* **50:** 1677-1686.

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

Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA *et al.* (2013). Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc Natl Acad Sci U S A* **110:** E488-E497.

Palenik B, Grimwood J, Aerts A, Rouze P, Salamov A, Putnam N *et al.* (2007). The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc Natl Acad Sci U S A* **104:** 7705-7710.

Pennington JT, Chavez FP (2000). Seasonal fluctuations of temperature, salinity, nitrate, chlorophyll and primary production at station H3/M1 over 1989-1996 in Monterey Bay, California. *Deep-Sea Res PT II* **47:** 947-974.

Rodriguez F, Derelle E, Guillou L, Le Gall F, Vaulot D, Moreau H (2005). Ecotype diversity in the marine picoeukaryote Ostreococcus (Chlorophyta, Prasinophyceae). *Environ Microbiol* **7:** 853-859.

Santoro A, Sakamoto C, Smith J, Plant J, Gehman A, Worden A *et al.* (2013). Measurements of nitrite production in and around the primary nitrite maximum in the central California Current. *Biogeosciences* **10:** 7395-7410.

Sherr EB, Sherr BF, Wheeler PA (2005). Distribution of coccoid cyanobacteria and small eukaryotic phytoplankton in the upwelling excosystem off the Oregon coast during 2001 and 2002. *Deep-Sea Res PT II* **52:** 317-330.

Simmons MP, Bachy C, Sudek S, van Baren MJ, Sudek L, Ares M, Jr. *et al.* (2015). Intron invasions trace algal speciation and reveal nearly identical Arctic and Antarctic *Micromonas* populations. *Mol Biol Evol* **32:** 2219-2235.

Simmons MP, Sudek S, Monier A, Limardo AJ, Jimenez V, Perle CR *et al.* (2016). Abundance and biogeography of picoprasinophyte ecotypes and other phytoplankton in the Eastern North Pacific Ocean. *Appl Environ Microbiol* **82:** 1693-1705.

Subirana L, Pequin B, Michely S, Escande ML, Meilland J, Derelle E *et al.* (2013). Morphology, genome plasticity, and phylogeny in the genus ostreococcus reveal a cryptic species, O. mediterraneus sp. nov. (Mamiellales, Mamiellophyceae). *Protist* **164:** 643-659.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol biol and evol* **30:** 2725-2729.

Treusch AH, Demir-Hilton E, Vergin KL, Worden AZ, Carlson CA, Donatz MG *et al.* (2012). Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. *ISME J* **6:** 481-492.

Tyrrell T (1999). The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature* **400:** 525-531.

van Baren MJ, Bachy C, Reistetter EN, Purvine SO, Grimwood J, Sudek S *et al.* (2016). Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC genomics* **17:** 1.

Vaulot D, Eikrem W, Viprey M, Moreau H (2008). The diversity of small eukaryotic phytoplankton (< or =3 microm) in marine ecosystems. *FEMS Microbiol Rev* **32:** 795-820.

Vaulot D, Lepere C, Toulza E, De la Iglesia R, Poulain J, Gaboyer F *et al.* (2012). Metagenomes of the picoalga Bathycoccus from the Chile coastal upwelling. *PLoS One* **7:** e39648.

Whitney FA (2015). Anomalous winter winds decrease 2014 transition zone productivity in the NE Pacific. *Geophys Res Let* **42:** 428-431.

Worden AZ, Not F (2008). Ecology and diversity of picoeukaryotes. In: Kirchman DL (ed). *Microbial Ecology of the Oceans*, 2nd edn. Wiley: Hoboken. p 594.

Worden AZ, Lee JH, Mock T, Rouze P, Simmons MP, Aerts AL *et al.* (2009). Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* **324:** 268-272.

Zaba KD, Rudnick DL (2016). The 2014–2015 warming anomaly in the Southern California Current System observed by underwater gliders. *Geophys Res Let* **43:** 1241-1248.

Zhu F, Massana R, Not F, Marie D, Vaulot D (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol* **52:** 79-92.

CHAPTER 3: Conclusions & Future Directions

Phylogenetic analysis of ITS sequences from environmental and cultured strains of *Bathycoccus* supported the notion of two distinct lineages. Previous studies have shown that divergence in a conserved region of the ITS2 (i.e. universal helix 3) is linked to sexual incompatibility in animals, plants and other protists (Coleman 2009). Given the limitations of using morphology and sexual compatibility based approaches to delineate species, this observation led to the proposal that ITS2 could serve as a molecular guide for identifying biological species in eukaryotes (Coleman 2009, Schultz and Wolf 2009, Yao et al 2010). Distinct *Ostreococcus* species that have been defined based on genome analyses have divergence in the ITS, however they are not different in the highly conserved ITS2 region that has been identified as important in other eukaryotes. Likewise, *Bathycoccus* clades BI and BII do not diverge in the conserved ITS2 region, however their full ITS shows a number of indels that would presumably correspond to sexual incompatibility and thus these ecotypes may correspond to different species. Future investigations could utilize comparative genomics to analyze ecotype-specific adaptations and further explore the assertion that *Bathycoccus* clades represent distinct species.

Referencing the *Bathycoccus* ITS sequences recovered in this study, I designed two qPCR assays to provide the first quantitative estimates of *Bathycoccus* clades BI and BII in a diverse set of samples from the North Pacific and tropical Atlantic, including two time-series studies. Additionally, *Ostreococcus* ecotypes were

enumerated using established qPCR assays for clades OI and OII. The OI and BI ecotypes were detected in cooler, mesotrophic ecosystems, and OII and BII ecotypes were present in warm, oligotrophic waters. The independent evolution of mesotrophic and oligotrophic ecotypes in these two genera suggests the same fundamental evolutionary pressures drive speciation in these closely-related picoeukaryotic green algae.

The concept of r- vs K-strategists is a useful approach for describing ecological niches in marine phytoplankton. *In-silico* ecosystem modelling studies have proposed that phytoplankton species can be split into r-strategist "bloomers" that possess a capacity for rapid growth when ideal conditions are present, and K-strategist "gleaners," which have low nutrient requirements and slow growth rates (Dutkiewicz et al 2009). The biogeographies of ecotypes observed in this study would suggest that mesotrophic clades BI and OI correspond to r-strategists that inhabit high latitude seas where strong seasonal mixing and variability favor phytoplankton with high growth rates. Oligotrophic clades BII and OII correspond to K-strategists and are found in tropical and subtropical latitudes where the water column is persistently stratified and nutrient fluxes to the photic zone are relatively low, thus selecting for taxa with lower cell quotas.

Bathycoccus and *Ostreococcus* ecotypes were observed co-existing in boundary regions, however both *Bathycoccus* ecotypes were observed at a wider range of environmental conditions and in more samples overall relative to *Ostreococcus* clades. Relative contributions from mesotrophic and oligotrophic clades in regions

where they co-occurred showed a strong correlation to depth, nutrients and salinity. Prior laboratory-based investigations have proposed that *Ostreococcus* clades OI and OII represent "high-light" and "low-light" adapted ecotypes, respectively (Cardol et al 2008, Rodriguez et al 2005), however field observations show that temperature and nutrient availability are more important parameters for explaining environmental distributions of *Ostreococcus* ecotypes (Demir-Hilton et al 2011, Simmons et al 2016, this study). Despite the observation that temperature and nutrients strongly influence biogeographic patterns of *Bathycoccus* ecotypes, the correlation between relative abundances and depth suggests *Bathycoccus* clades BI and BII could be adapted to distinct light environments, therefore follow-up studies could explore differences in photosynthetic physiology between *Bathycoccus* clades BI and BII.

The 18S rRNA gene is commonly used to assess eukaryotic biodiversity and to estimate abundances of different phytoplankton using methods such as qPCR and FISH. Importantly, the 18S V9 region does not allow discrimination of the two *Ostreococcus* ecotypes (Monier et al 2016), and the *Bathycoccus* ecotypes do not exhibit any divergence in the 18S rRNA gene. Thus, investigations based on these gene markers do not resolve ecotypes of picoeukaryotes that we show have distinct biogeographies and appear to partition at the species level. These results suggest that higher resolution gene markers, such as the ITS, are necessary to discriminate environmentally relevant levels of diversity within the picoeukaryotic size class. Furthermore, investigations with larger phytoplankton (i.e. diatoms and dinoflagellates) have shown that ITS sequences are extremely successful at

delineating biological species, and often outperformed rRNA gene markers (Amato et al 2007, Litaker et al 2007, Moniz and Kaczmarska 2010), therefore this approach would likely be beneficial for surveys of diversity that target many different taxa of phytoplankton.

The clade-specific qPCR assays utilized in this study were able to detect rapid shifts in the picoeukaryotic community of Monterey Bay that were connected to the arrival of an anomalously warm water mass, demonstrating the potential for these phytoplankton to act as indicator species that reflect changing oceanographic conditions. Furthermore, increased thermal stratification associated with climate change is predicted reduce fluxes of nutrients and expand oligotrophic regions in oceans around the globe (Behrenfeld et al 2006), thus causing oligotrophic ecotypes to displace mesotrophic phytoplankton in certain habitats. Currently, it is unknown how changes in picoeukaryotic community structure impact primary productivity and energy transfer to higher trophic levels. Based on the findings presented here, subsequent investigations could utilize ecotypes of *Bathycoccus* and *Ostreococcus* as environmentally relevant model organisms to explore how marine ecosystems will respond to future habitat disturbances.

References

Amato A, Kooistra WH, Ghiron JHL, Mann DG, Pröschold T, Montresor M (2007). Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist* **158:** 193-207.

Behrenfeld MJ, O'Malley RT, Siegel DA, McClain CR, Sarmiento JL, Feldman GC *et al* (2006). Climate-driven trends in contemporary ocean productivity. *Nature* **444:** 752-755.

Cardol P, Bailleul B, Rappaport F, Derelle E, Beal D, Breyton C *et al* (2008). An original adaptation of photosynthesis in the marine green alga Ostreococcus. *Proc Natl Acad Sci U S A* **105:** 7881-7886.

Coleman AW (2009). Is there a molecular key to the level of "biological species" in eukaryotes? A DNA guide. *Mol Phylogenet Evol* **50:** 197-203.

Demir-Hilton E, Sudek S, Cuvelier ML, Gentemann CL, Zehr JP, Worden AZ (2011). Global distribution patterns of distinct clades of the photosynthetic picoeukaryote Ostreococcus. *ISME J* **5:** 1095-1107.

Dutkiewicz S, Follows MJ, Bragg JG (2009). Modeling the coupling of ocean ecology and biogeochemistry. *Global Biogeochem Cycles* **23**(4).

Litaker WR, Vandersea MW, Kibler SR, Reece KS, Stokes NA, Lutzoni FM *et al* (2007). Recognizing dinoflagellate species using ITS rDNA sequences. *J Phycol* **43:** 344-355.

Monier A, Worden AZ, Richards TA (2016). Phylogenetic diversity and biogeography of the Mamiellophyceae lineage of eukaryotic phytoplankton across the oceans. *Environ Microbiol Rep* **8**(4): 461-469.

Moniz MB, Kaczmarska I (2010). Barcoding of diatoms: nuclear encoded ITS revisited. *Protist* **161:** 7-34.

Rodriguez F, Derelle E, Guillou L, Le Gall F, Vaulot D, Moreau H (2005). Ecotype diversity in the marine picoeukaryote Ostreococcus (Chlorophyta, Prasinophyceae). *Environ Microbiol* **7:** 853-859.

Schultz J, Wolf M (2009). ITS2 sequence–structure analysis in phylogenetics: a how-to manual for molecular systematics. *Mol Phylogenet Evol* **52:** 520-523.

Yao H, Song J, Liu C, Luo K, Han J, Li Y *et al* (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *PloS one* **5:** e13102.