Quantitative biogeography of picoprasinophytes establishes ecotype distributions and significant contributions to marine phytoplankton

Alexander J. Limardo,1,2 Sebastian Sudek,1* Chang Jae Choi,1 Camille Poirier,1 Yoshimi M. Rii,3 Marguerite Blum,1 Robyn Roth,4 Ursula Goodenough,5 Matthew J. Church3 and Alexandra Z. Worden1,2**

1Monterey Bay Aquarium Research Institute, Moss Landing, CA, USA.
2University of California Santa Cruz, Santa Cruz, CA, USA.
3University of Hawaii at Manoa, Honolulu, HI, USA.
4Washington University School of Medicine, St. Louis, MO, USA.
5Washington University, St. Louis, MO, USA.

Summary

Bathycoccus and Ostreococcus are broadly distributed marine picoprasinophyta algae. We enumerated small phytoplankton using flow cytometry and qPCR assays for phylogenetically distinct Bathycoccus clades BI and BII and Ostreococcus clades OI and OII. Among 259 photic-zone samples from transects and time-series, Ostreococcus maxima occurred in the North Pacific coastal upwelling for OI (36 713 ± 1485 copies ml⁻¹) and the Kuroshio Front for OII (50 189 ± 561 copies ml⁻¹) and the two overlapped only in frontal regions. The Bathycoccus overlapped more often with maxima along Line-P for BI (10 667 ± 1299 copies ml⁻¹) and the tropical Atlantic for BII (4125 ± 339 copies ml⁻¹). Only BII and OII were detected at warm oligotrophic sites, accounting for 34 ± 13% of 1589 ± 448 eukaryotic phytoplankton cells ml⁻¹ (annual average) at Station ALOHA’s deep chlorophyll maximum. Significant distributional and molecular differences lead us to propose that Bathycoccus clade BII represents a separate species which tolerates higher temperature oceanic conditions than Bathycoccus prasinos (BI).

Introduction

Prasinophytes are green algae that are present in many marine environments. Members of the class Mamiellophyceae are particularly widespread and include both picoplanktonic (<2 μm diameter) and larger genera. Early 18S rRNA gene clone libraries from the picoplankton size class contained numerous sequences from the Mamiellophyceae Bathycoccus, Micromonas and Ostreococcus (Guillou et al., 2004; Worden, 2006; Massana, 2011). Mamiellophyceae were also present at all 47 sites sampled during Tara Oceans based on 18S rRNA V9 amplicon data (de Vargas et al., 2015) and comprised a large fraction of photic-zone amplicons in two polar studies (Monier et al., 2014; Metfies et al., 2016). Cell abundances of these Mamiellophyceae taxa have been compared by FISH and qPCR in mesotrophic or coastal (Not et al., 2005; Countway and Caron, 2006; Collado-Fabbri et al., 2011) and oligotrophic (Zhu et al., 2005; Not et al., 2008; Treusch et al., 2012; Gomez-Pereira et al., 2013) environments. High abundance has been documented in more nutrient rich environments, and in oligotrophic systems peak abundances appear to be related to periods when nutrient levels are higher—for example, springtime in the Sargasso Sea—and then decline to low or undetectable numbers when the water column is strongly stratified (Treusch et al., 2012). A summary of this literature indicates that Ostreococcus is restricted to temperate and tropical waters, while Bathycoccus and Micromonas are present from low to high latitudes, including polar seas.

Understanding the dynamics of Bathycoccus, Micromonas and Ostreococcus is complicated by the fact that each genus contains several species and/or phylogenetically distinct clades. Comparative genomics and evolutionary
analyses underscore the differences between clades (Moreau et al., 2012; van Baren et al., 2016). The Micromonas genus exhibits the most genetic diversity and has seven distinct clades, two of which have genome-sequenced representatives (van Baren et al., 2016). Ostreococcus has four clades based on the 18S rRNA gene, all of which contain cultured representatives with sequenced genomes: Ostreococcus lucimarinus (Ostreococcus clade A), Ostreococcus sp. RCC809 (clade B), Ostreococcus tauri (clade C) and Ostreococcus mediterraneus (clade D) (Guillou et al., 2004; Subirana et al., 2013). Bathycoccus and Ostreococcus are evolutionarily closer to each other than to Micromonas, and both are non-motile, unlike other prasinophytes including Micromonas. Bathycoccus is more akin to most other prasinophytes in having scales, which both Ostreococcus and Micromonas lack. All Bathycoccus cultures and environmental sequences obtained thus far have identical 18S rRNA genes, unlike the various Micromonas and Ostreococcus clades (Guillou et al., 2004; Monier et al., 2013).

Despite initial consideration as a single species, Bathycoccus prasinos, flow cytometric targeted metagenome studies demonstrated diversity in ITS2–5.8S sequences and other genes, leading to the hypothesis that at least two Bathycoccus ecotypes exist (Monier et al., 2012; Vaulot et al., 2012; Monier et al., 2013). Subsequent analyses of single-copy homologous genes (ecomarkers) in targeted metagenomic/metatranscriptomic data and the sequenced genome of B. prasinos strain RCC1105 (Moreau et al., 2012) showed that two Bathycoccus types exist with 82 ± 6% nucleotide identity across homologs (Simmons et al., 2016). In contrast, homologs from Ostreococcus clade OI and OII ecotypes, corresponding to Ostreococcus clades A and B, respectively, displayed 75 ± 8% nucleotide identity. Likewise, comparison of a more recent targeted metagenome assembly of a wild Bathycoccus population and the RCC1105 genome also led to the conclusion that two types exist (Vannier et al., 2016). These two recent omic-based studies also mapped reads to the respective wild and cultured genome assemblies, leading to the suggestion that the two Bathycoccus clades, BI and BII, have largely non-overlapping distributions (Simmons et al., 2016; Vannier et al., 2016) akin to patterns established previously for Ostreococcus OI and OII using more quantitative methods, specifically qPCR (Demir-Hilton et al., 2011). 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) (Demir-Hilton et al., 2011). Additionally, the OI and OII clades co-occurred only at frontal locations where physical mixing occurs between more coastal and oligotrophic waters. Thus, the Ostreococcus clades represent ecotypes adapted to mesotrophic (OI) and oligotrophic (OII) conditions (Demir-Hilton et al., 2011; Clayton et al., 2017).

We set out to test whether the proposed BI and BII groups exhibit abundance distributions like Ostreococcus, where the OI and OII clades rarely overlap except in dynamic frontal regions (Demir-Hilton et al., 2011; Simmons et al., 2016; Clayton et al., 2017). To this end, we developed qPCR assays that resolve Bathycoccus ITS clades that have identical 18S rRNA gene sequences. Clone libraries that recovered the ITS1/5.8S/ITS2 regions of the rRNA transcriptional unit were constructed and sequenced from four sites and three cultured strains to improve phylogenetic resolution. We then quantified and compared BI, BII, OI and OII in samples from several regions of the North Pacific Ocean, the tropical Atlantic and two time-series sites, Station ALOHA near Hawaii and Station M1 in Monterey Bay. The abundance and distributions of these four Mamiellophyceae groups were analysed in the context of available environmental data and eukaryotic phytoplankton cell abundances enumerated by flow cytometry. The results demonstrate that clades BI and BII overlap more frequently than the two Ostreococcus ecotypes. However, the two Bathycoccus clades represent ecotypes with distinct biogeographies and have molecular differences that suggest clade BII is a separate species from B. prasinos. Moreover, our quantitative cell abundance data indicates that recent 18S rRNA V9 amplicon-based studies underestimate the importance of these taxa in the open-ocean.

Results

Bathycoccus and Ostreococcus ITS phylogeny

Environmental ITS1/5.8S/ITS2 clone libraries rendered Bathycoccus sequences from the tropical Atlantic, tropical Pacific, Kuroshio Front and Monterey Bay, and Ostreococcus sequences from all but Monterey Bay (Fig. 1A, Table 1, Supporting Information Table S1). Sequences were also recovered from Micromonas (Supporting Information Table S1) and other algae. Our maximum-likelihood phylogenetic analyses resolved two Bathycoccus clades (100% bootstrap support; Fig. 1B). The overall topology was similar to a prior reconstruction for Mamiellophyceae (Marin and Melkonian, 2010), except the latter had just one Bathycoccus clade. In that study, only cultured taxa with deposited sequences were included, whereas we sequenced and analysed environmental clone libraries as well as isolates from the Indian Ocean (RCC715 and RCC716) expanding the repertoire of prasinophyte ITS sequences available at that time. Monterey Bay Bathycoccus ITS sequences had on average 99% nucleotide identity to B. prasinos RCC1105, a coastal Mediterranean isolate. The Indian Ocean isolates and most sequences from the other sites belonged to a separate clade, for which average ITS
nucleotide identity to RCC1105 was 95%. The Bathycoccus clade containing B. prasinos RCC1105 is hereafter referred to as ‘clade BII’ and the other clade as ‘clade BII’, according to a previous proposal (Simmons et al., 2016).

In addition to phylogenetic differences, we observed structural differences between the BI and BII ITS sequences (Supporting Information Fig. S1a, b). For example, in the ITS2 of BII sequences universal Helix 1 was extended by an extra 7 base pairs, comprised of 2 G-C and 5 A-U from position 635 to 651 of environmental clone TAM24 (KY368638), along with a minimal 3 nt loop (Supporting Information Fig. S1c), following base pair 8 of B. prasinos (e.g., CCMP1898, RCC1105 and BI environmental clones) Helix 1. Helix 2 also exhibited differences as did the additional helix specific to the family Bathycoccaceae (Marin and Melkonian, 2010) between universal Helices 3 and 4 (Supporting Information Fig. S1d), as well as universal Helix 4. Differences between BI and BII and from other members of the class Mamiellaphyceae were also observed in the ITS1 (e.g., Supporting Information Fig. S2).

**Electron microscopy of Bathycoccus and Ostreococcus clades**

Given the distinctive clades of Bathycoccus and Ostreococcus and the fact that ultrastructural information was not available for BII and OII isolates, we used Quick-Freeze Deep-Etch (QFDE), thin-section and negative-stain electron microscopy to look for characteristic traits and potential distinguishing features. Our phylogenetic analyses demonstrated that RCC715 and RCC716 were members of the BI clade (Fig. 1B), although earlier this clade was thought to be uncultured, providing candidate strains for BI imaging. The cytoplasmic organization and organelle morphology of the four clades (OI, OII, BI and BII) proved to be indistinguishable (Fig. 2A–D). A single mitochondrion, only infrequently displaying cristae, lies between the basal cup-shaped chloroplast and the anterior sausage-shaped nucleus, as evident in prior thin-sectioned preparations of B. prasinos, O. tauri and O. lucimarinus (Eikrem and Throndsen, 1990; Chretiennot et al., 1995; Worden et al., 2004). Each BI and BII nucleus carries 2–3 closely spaced nuclear pores (e.g., Fig. 2E) that have not been reported previously in these taxa apart from O. tauri (Henderson et al., 2007). The stacked thylakoid membranes displayed, as expected but not also previously reported in picoprasinophytes, large intramembranous particles associated with photosystem II (Goodenough and Staehelin, 1971) (Fig. 2H). One notable difference was that the etched outer surface of the OII cell membrane carried knob-like protuberances absent from the other three taxa (Supporting Information Fig. S3a-c).

The scales of both Bathycoccus clades form overlapping layers on the cell surface (Fig. 2I and Supporting Information Fig. S3d). Scale size and organization were indistinguishable between BI and BII (Fig. 2K and L and Supporting Information Fig. S3g). Images of the Golgi apparatus (Fig. 2F and G) showed no evidence of scales in the cisternae or vesicles, nor were secretion pores present. Instead, embossed scales (boxed, Fig. 2I) were observed beneath the cell membrane, as was a naked scale that was exposed, via fracture, just beneath the cytoplasm-facing P-face of the cell membrane (boxed,
Fig. 2J). In flattened negatively stained images, the eight projections from the central hub of the scale are coplanar, but in QFDE replicas, the scales often adopt a cupped configuration (Supporting Information Fig. S3g and h) such that four of the eight are more prominent, generating an X-shape (Fig. 2K and Supporting Information Fig. S3d).

Temporal dynamics of Bathycoccus and Ostreococcus clades by qPCR

We investigated temporal changes in the abundance of the Bathycoccus and Ostreococcus clades at Station M1 in Monterey Bay and Station ALOHA in the North Pacific Subtropical Gyre. Primer-probe sets to individually quantify clade BI and clade BII were designed based on the observed differences between environmental and culture ITS sequences and specifically targeted the ITS1 (Supporting Information Tables S2 and S3 and Figs S2 and S4). These were implemented alongside established Ostreococcus OI and OII 18S rRNA gene qPCR assays (Demin-Hilton et al., 2011) and hereafter data from the Bathycoccus and Ostreococcus primer-probe sets is provided as rDNA copies ml\(^{-1}\) for simplicity.

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 NO\(_3\), NO\(_2\), PO\(_4\)\(^3-\) and SiO\(_4\)\(^2-\) were maximal, indicating the influence of wind-driven coastal upwelling (Fig. 3, Supporting information). Thermal stratification increased during summer as surface temperatures increased from 13.6°C (10 July) to a peak of 17.3°C (22 September). Longer-term observations show surface temperatures rarely exceeding 15°C during late summer and autumn (Pennington and Chavez, 2000), but our study included an anomalously warm patch of water (Bond et al., 2015). This resulted in abnormally low nutrient concentrations, where NO\(_3\) was low (<1.0 μM) from 12 August until the study end, and PO\(_4\)\(^3-\) declined from a maximum (1.4 μM) on 5 May to a minimum (0.2 μM) on 29 October (Supporting information). Decreased salinity between July and late October reflected 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. 3). The two depths analysed, surface (0–5 m) and the base of the wind-mixed surface layer (~10–20 m), and chlorophyll maximum if present, trended together over time. Clade BI was detected throughout the study and was most abundant in October and November, peaking at 8444 ± 529 rDNA copies ml\(^{-1}\) (20 m, 20 November). Clade BII was only detected between August and November, with lower overall abundance than BI (p < 0.0001) and a peak of 775 ± 71 rDNA copies ml\(^{-1}\) (20 m, 20 November). Clade OI was also present throughout the entire study period, but showed higher abundances in spring and summer, with a maximum of 18 033 ± 1260 rDNA copies ml\(^{-1}\) (19 June). Clade OII was only detected between August and November and did not exceed 50 rDNA copies ml\(^{-1}\).

Temperature was significantly higher (15.7°C ± 1.0°C, n = 7) in samples where BI was detected, than where it was absent (13.4°C ± 2.0°C), and salinity was lower in BI containing samples due to intruding California Current waters (p < 0.05). Additionally, NH\(_4\)\(^+\) (0.06 ± 0.08 μM), NO\(_3\)\(^-\) (1.22 ± 2.29 μM) and N:P ratios of the dissolved pool (1.35 ± 1.88) were lower (p < 0.05) than in samples where BI was absent (NH\(_4\)\(^+\), 0.25 ± 0.25 μM; NO\(_3\)\(^-\), 4.57 ± 3.26 μM; and N:P ratio 4.59 ± 2.56). Results were similar for OII, which was present in 7 samples, 6 of which also contained BII (Fig. 3).

Station ALOHA samples were collected on 10 dates between July 2011 and August 2012. Over this period, strong thermal stratification in the photic-zone (<200 m) weakened only in the winter months (Fig. 4). Sea surface temperatures varied from a maximum of 26.3°C (September 2011) to a minimum of 23.0°C (March 2012). The DCM was located between 100 m and 150 m in depth.

Table 1. Cruises analysed, including 10 HOT cruises to Station ALOHA and 11 MBTS cruises to Station M1.

| Location                           | Cruise          | Station(s) | Start (or station) | End point | Dates       |
|------------------------------------|-----------------|------------|--------------------|-----------|-------------|
| N. Pacific Subtropical Gyre       | HOT             | ALOHA      | Lat 12.75 Lon -158.00 | na        | 06/11–08/12 |
| MB                                 | MBTS            | M1         | Lat 36.75 Lon -122.00 | na        | 03/14–11/14 |
| East. N. Pacific (Line-67)        | WFD09           | 9          | Lat 36.80 Lon -121.85 | 33.29     | 10/09       |
| East. N. Pacific                   | CANON11         | 3          | Lat 36.13 Lon -123.49 | 33.95     | 09/11       |
| East. N. Pacific                   | CANON10         | 2          | Lat 35.92 Lon -122.90 | 35.79     | 09/10       |
| East. N. Pacific                   | MV1405          | 3          | Lat 38.24 Lon -126.62 | 35.93     | 07/14       |
| MB to Gulf of California           | GOC12           | 9          | Lat 36.70 Lon -126.38 | 21.00     | 02/12       |
| Sub-Arctic Pacific (Line-P)       | GeOMICS         | 4          | Lat 48.58 Lon -125.50 | 48.82     | 05/12       |
| Kuroshio Front                     | NT09–18         | 25         | Lat 36.61 Lon 143.48 | 35.17     | 10/09       |
| Tropical Atlantic                  | SJ0609          | 8          | Lat 11.67 Lon -51.49 | 1.83      | 06/06–07/06 |

na, not applicable; MB, Monterey Bay.
Average NO$_3^-$ (0.02 ± 0.01 μM) and PO$_4^{3-}$ (0.09 ± 0.05 μM) concentrations were low at the surface (5 m, Supporting Information). Neither BI nor OI were detected at ALOHA, whereas BII and OII were present year-round. Clades BII and OII were present at low abundance at depths ≤50 m from November to May and were below detection limits at these depths during warmer, strongly stratified periods (Fig. 4). Maximum abundance tended to coincide with the depth of the DCM. Peak abundances were comparable, with 1146 ± 135 rDNA copies ml$^{-1}$ (BII).

© 2017 The Authors. Environmental Microbiology published by Society for Applied Microbiology and John Wiley & Sons Ltd., Environmental Microbiology, 19, 3219–3234

---

**Fig. 2.** Ultrastructure of cultured representatives of the two Bathycocccus and two Ostreococcus clades by QFDEEM. Survey cross-fracture views of A. clade BI, B. clade BII, C. clade OI and D. clade OII. Higher-magnification views of BII show E. three nuclear pores, F. G. the Golgi apparatus and H. chloroplast. In the latter, large intramembranous particles containing photosystem II (arrow), and cross-fractured single thylakoids (*) are visible. Scale bars of A–H: 100 nm. I., J. Scale system of clade BII. I. P-fracture face of cell membrane in center, surrounded by scales exposed by etching. Embossed scales beneath the membrane are boxed. Scale bar: 250 nm. J. P-fracture face of cell membrane; in boxed area, fracture has dipped down into cytoplasm to expose a naked scale. Scale bar: 100 nm. K. Negatively stained and L. deep-etched scale(s), scale bars: 50 nm. The following strains were imaged: *B. prasinos* CCMP1898 (BI), RCC716 (BII), *O. lucimarinus* (OI) and O. RCC809 (OII). Abbreviations: c, chloroplast; er, endoplasmic reticulum; g, Golgi apparatus; m, mitochondrion; mi, microbody; n, nucleoplasm; ne, nuclear envelope; st, starch.
ward along the North American coast to tropical waters, beginning in Monterey Bay and transiting south through the Kuroshio Front region and the tropical Atlantic. We examined samples from a transect from coastal Monterey Bay southwestward, crossing the core of the California Current to 1000 km offshore, while OII was reported at the two most oligotrophic stations in the DCM (maximum abundance 374,000 rDNA copies ml−1 at 75 m) and NO3− concentrations of 4.48 μM and NO4− (1.20 ± 2.31 μM) as well as N:P ratios (1.30 ± 1.92) were lower than samples where OII was present compared with depths where OII was not detected (>0.01). This suggests that their persistence deep in the photic-zone (Fig. 4) is driven by nutrient availability, particularly nitrogen or co-associated variables.

Biogeography of Bathycoccus and Ostreococcus clades

To broaden results from the temporal studies at Station M1 and Station ALOHA, we examined samples from a transect cruise (WFAD09) along the California Cooperative Oceanic Fisheries Investigations Line-67, a cruise (GOC12) beginning in Monterey Bay and transiting southward along the North American coast to tropical waters, and additional samples from cruises along Line-67, Line-P, in the Kuroshio Front region and the tropical Atlantic (Fig. 1, Table 1, Supporting Information). Along the Line-67 transect from coastal Monterey Bay southwestward, crossing the core of the California Current to ~1000 km offshore at the edge of oligotrophic North Pacific Subtropical Gyre, BI decreased moving westward. BI abundances were comparable between the transition zone (stations 67–65 and 67–95) and oligotrophic (stations 67–145 and 67–155) stations and this clade was not detected only in Monterey Bay (~5 km from shore, Fig. 5). Clades OI and OII overlapped less frequently than BI and BII. OI was present at a maximum of 14,027 rDNA copies ml−1 near the coast and in the transition zone, but was not detected farther offshore, while OII was reported at the two most oligotrophic stations in the DCM (maximum abundance 72 ± 3 rDNA copies ml−1, 100 m depth) but otherwise was found only at the westward edge of the transition zone (station 67–95) (Simmons et al., 2016).

The GOC12 cruise transected the cool, upwelling-influenced waters of Monterey Bay through warmer waters (up to 24.2°C) 200 km south of Baja California, Mexico. Nutrient concentrations were upwards of 4.48 μM NO3− and 0.61 μM PO4− in the near surface waters (<5 m) of Monterey Bay, and decreased to below detection (<0.01 μM) versus not detected (0.09 μM) as thermal stratification strengthened in spring (30 May 2012). PO4− concentrations were not significantly different in samples where BI and OII were present (0.1 ± 0.04 μM) versus not detected (0.09 ± 0.04 μM). However, NO3− concentrations were significantly higher where these taxa were present compared with depths where they were not detected (0.36 ± 0.54 and 0.02 ± 0.01 μM, respectively, p < 0.01, Supporting Information Table S4). This suggests that their persistence deep in the photic-zone (Fig. 4) is driven by nutrient availability, particularly nitrogen or co-associated variables.

Fig. 3. Dynamics of Ostreococcus and Bathycoccus and clades in Monterey Bay, California. Samples were collected at MBTS Station M1 (Pennington and Chavez, 2000). A. Contour plot of PO4− over the study period. Black dots represent CTD data point locations used for extrapolation. B. Contour plot of temperature. C. qPCR abundance 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 the subsurface chlorophyll max, if present, or the base of the wind-mixed surface layer. D. Bathycoccus clade BI and Ostreococcus clade OII qPCR abundances over the study period. Like BI, OII was also found where temperatures were higher (15.8°C ± 1.1°C) and concentrations of NH4+ (0.07 ± 0.07 μM) and NO3− (1.20 ± 2.31 μM) as well as N:P ratios (1.30 ± 1.92) were lower than samples where OII was not detected. Note the different Y-axis scales in panel C and C.

Biogeography of Bathycoccus and Ostreococcus clades

To broaden results from the temporal studies at Station M1 and Station ALOHA, we examined samples from a transect cruise (WFAD09) along the California Cooperative Oceanic Fisheries Investigations Line-67, a cruise (GOC12) beginning in Monterey Bay and transiting southward along the North American coast to tropical waters, and additional samples from cruises along Line-67, Line-P, in the Kuroshio Front region and the tropical Atlantic (Fig. 1, Table 1, Supporting Information). Along the Line-67 transect from coastal Monterey Bay southwestward, crossing the core of the California Current to ~1000 km offshore at the edge of oligotrophic North Pacific Subtropical Gyre, BI decreased moving westward. BI abundances were comparable between the transition zone (stations 67–65 and 67–95) and oligotrophic (stations 67–145 and 67–155) stations and this clade was not detected only in Monterey Bay (~5 km from shore, Fig. 5). Clades OI and OII overlapped less frequently than BI and BII. OI was present at a maximum of 14,027 ± 548 rDNA copies ml−1 near the coast and in the transition zone, but was not detected farther offshore, while OII was reported at the two most oligotrophic stations in the DCM (maximum abundance 72 ± 3 rDNA copies ml−1, 100 m depth) but otherwise was found only at the westward edge of the transition zone (station 67–95) (Simmons et al., 2016).

The GOC12 cruise transected the cool, upwelling-influenced waters of Monterey Bay through warmer waters (up to 24.2°C) 200 km south of Baja California, Mexico. Nutrient concentrations were upwards of 4.48 μM NO3− and 0.61 μM PO4− in the near surface waters (<5 m) of Monterey Bay, and decreased to below detection (<0.01 μM) versus not detected (0.09 μM) as thermal stratification strengthened in spring (30 May 2012). PO4− concentrations were not significantly different in samples where BI and OII were present (0.1 ± 0.04 μM) versus not detected (0.09 ± 0.04 μM). However, NO3− concentrations were significantly higher where these taxa were present compared with depths where they were not detected (0.36 ± 0.54 and 0.02 ± 0.01 μM, respectively, p < 0.01, Supporting Information Table S4). This suggests that their persistence deep in the photic-zone (Fig. 4) is driven by nutrient availability, particularly nitrogen or co-associated variables.
Both picoprasinophyte genera were detected at every station sampled, and the four clades co-occurred over a large region (Supporting Information). How-

ever, BI and OI were not detected at the oligotrophic southernmost station, and BII and OII were not detected at the two northernmost stations closer to Monterey Bay. Ratios of BI:BII, or OI:OII, in WFAD09 and GOC12 samples where both clades co-occurred (53 out of 74 total) trended negatively with depth and phosphate concentra-
tions (Supporting Information Fig. S6).

In other samples, clade abundances varied in a manner consistent with results from the time-series studies and transects above. Clade BII and clade OII were not detected along Line-P, while BI and OI abundances increased from 803 and 625 rDNA copies ml⁻¹ at coastal Station P1 to 10 667 and 35 797 rDNA copies ml⁻¹ at offshore Station P6, respectively. BI abundance declined at the most westward station (P8, 1405 ± 73 rDNA copies ml⁻¹) where OI abundances were an order of mag-
nitude greater (13 766 ± 1843 rDNA copies ml⁻¹) than BI. Cruises CANON11, S410 and MV1405 (Table 1) were dominated by BI and OI, while BII and OII were detected at the southern most station sampled during MV1405, in the region of the anomalously warm patch. In the Kur-
oshio/Oyashio Frontal zone both Bathycoccus clades were found, with maxima of 1751 rDNA copies ml⁻¹ (BI) and 2662 rDNA copies ml⁻¹ (BII). High OII abundances (50 189 rDNA copies ml⁻¹) occur here, while OI was also present but less abundant (Clayton et al., 2017). Lastly, OII was previously detected in one tropical Atlantic sample, at 73 rDNA copies ml⁻¹ (Demir-Hilton et al., 2011), while clade BII was present in all the tropical Atlantic samples analysed here. The BII maximum (4125 ± 339 rDNA copies ml⁻¹) was in a sample where NO₃ was 0.03 μM and PO₄ was below detection limits (<0.01 μM, 61 m depth), and neither OI nor BI were detected. These results illustrate a general trend: OI and BI are present in cooler, lower salinity waters that ranged here from 7.8°C to 24.1°C (OI) or 25.2°C (BII), while OII and BII are present in warmer, higher salinity waters that ranged in our samples from 11.3°C to 27.3°C (OII) and 9.9°C to 28.9°C (BII) (Fig. 6, Supporting Information). Additionally, both OII and BII were found in highly stratified systems and peak abundan-
ces appeared to track the nutricline.

Statistical analyses using data from across multiple regions

Additional statistical tests using qPCR and environmental data examined trends in the 230 samples where at least one clade of Bathycoccus or Ostreococcus was detected.
and the metadata matrix criteria were met. The complete sample set was skewed towards surface and coastal samples and, given the breadth of environments sampled, associations with depth also reflect differences in water column structure, irradiance, nutrient availability and other factors. Additionally, NO$_3^-$, NO$_2^-$ and NH$_4^+$ were below detection in multiple open-ocean surface samples, 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 BI and OI were negatively correlated with inorganic nutrients (for PO$_4^{3-}$, $\rho = -0.30$ and $\rho = -0.44$, respectively) due to their more frequent presence in oligotrophic habitats. Significant positive correlations were observed for BI and depth (Spearman’s correlation $\rho = 0.25$) and between OI and chlorophyll $a$ concentrations ($\rho = 0.49$). Temperature and salinity were negatively correlated with abundance of BI ($\rho = -0.65$ and $\rho = -0.76$) and 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. 6, Supporting Information Table S5). In contrast, abundances of BI ($\rho = 0.43$ and $\rho = 0.42$) and OI ($\rho = 0.57$ and $\rho = 0.61$) were positively correlated with temperature and salinity, reflecting their distributions in tropical, subtropical and warmer temperate environments (Fig. 6, Supporting Information Table S5).

Multivariate transformation of physicochemical parameters was also combined with a clustering analysis to define habitat types. K-means clustering partitioned samples into three groups (Supporting Information Fig. S7) which were then used to bin qPCR abundance data and further describe the biogeography of each clade. Cluster 1 was dominated by clades BI and OI and contained samples from the North Pacific Subtropical Gyre and tropical Atlantic (latitudinal range: $2^\circ$ to $36^\circ$N), typically characterized by warm ($23.1 ^\circ\text{C} \pm 2.0 ^\circ\text{C}$), nutrient-poor ($0.51 \pm 0.67$ μM NO$_3^-$ and $0.11 \pm 0.09$ μM PO$_4^{3-}$), stratified, deep photic-zone environments like Station ALOHA (Supporting Information Fig. S7). In addition, samples collected from the Kuroshio/Oyashio frontal region also grouped with Cluster 1. BI and OI were not detected in Cluster 1 except for some Kuroshio samples. Clusters 2 and 3 contained higher abundances of BI and OI than Cluster 1. They also partitioned temperate samples (latitudinal range: $21^\circ$ to $49^\circ$N) into regions where the concentration of previously abbreviated as NO$_3^-$ and the N:P ratio were relatively higher (Cluster 3) or lower (Cluster 2). Among the three clusters, Cluster 2 had the highest median abundances of both BI and OI (Supporting Information Fig. S7).

**Contributions of Bathycoccus and Ostreococcus to eukaryotic phytoplankton**

Flow cytometry data were analysed for 190 samples from below $37^\circ$N (none were available from higher latitude samples). Eukaryotic phytoplankton enumerated by flow cytometry averaged from 1000 ± 245 cells ml$^{-1}$ annually in ALOHA surface samples to 20 983 ± 6038 ml$^{-1}$ in mesotrophic water columns (Table 2). Our analyses of genome sequences from the most environmentally-relevant Mamiellophyceae (*B. prasinos*, *O. lucimarinus*, *O* sp. RCC809, *M. pusilla*, *M. commoda*) showed that there
are two copies of the rRNA transcriptional units. Therefore, we presumed that qPCR derived abundances reflect twice the cellular abundances of the taxa studied. Contributions of *Bathycoccus* and *Ostreococcus* varied substantially from site to site, but overall averages were similar (Table 2). Together these two genera comprised around 20% of eukaryotic cells in the Monterey Bay samples and mesotrophic waters. At ALOHA, surface contributions were minimal but they comprised 34 ± 13% of the DCM eukaryotic phytoplankton, and 24 ± 17% throughout the water column during winter mixing (Table 2).

**Discussion**

Distinct niches of the *Ostreococcus* OI and OII ecotypes have been established using quantitative methods for enumeration in the field (Demir-Hilton et al., 2011; Simmons et al., 2016; Clayton et al., 2017), whereas the idea that

**Table 2.** Environmental parameters and average phytoplankton abundances by flow cytometry and qPCR (*Bathycoccus* plus *Ostreococcus* only, corrected for rDNA copy numbers) in samples from below latitude 37°N.

| Parameter                      | ALOHA          | Multiple Stratified Open Ocean | MB                        | Meso-trophic  |
|--------------------------------|----------------|--------------------------------|---------------------------|---------------|
|                                | Surf | DCM  | Winter | Surf | DCM  | Winter | Column | Column |
| **NO**₃ (µM)                   | 0.02 (0.01) | 0.16 (0.10) | 0.18 (0.31) | 0.02 (0.03) | 0.14 (0.25) | 4.97 (3.53) | 3.06 (2.68) |
| **PO**₄ (µM)                   | 0.09 (0.05) | 0.09 (0.02) | 0.10 (0.02) | 0.26 (0.20) | 0.15 (0.21) | 0.72 (0.22) | 0.61 (0.28) |
| **Temp (°C)**                  | 24.96 (1.06) | 22.29 (0.56) | 23.04 (0.99) | 22.95 (3.84) | 22 (6) | 12.90 (0.95) | 15.73 (3.41) |
| **Eukaryotic Phyto. (cells ml⁻¹)** | 1000 (245) | 1589 (448) | 833 (344) | 1311 (364) | 4072 (2480) | 20 983 (6038) | 16 872 (8193) |
| **Bathy+Ostreo (cells ml⁻¹)**  | 9 (26) | 538 (241) | 178 (123) | 7 (6) | 831 (658) | 3266 (1584) | 2808 (1622) |
| **%Bathy+Ostreo**              | 1 (2) | 34 (13) | 24 (17) | 1 (1) | 21 (9) | 16 (8) | 20 (12) |

The DCM average depth over the ALOHA time-series was 125 ± 18 m; ’Winter’ reflects vertical averages during deeper mixing at ALOHA; ’Multiple Stratified Open Ocean’ combines data from oligotrophic Line 67, Tropical Atlantic and a weighted ALOHA average (accounting for strongly stratified months versus less stratified); MB, Monterey Bay includes averaged data from M1 and M2 from 2009, 2012 and MBTS data (2014–2015); Mesotrophic samples are those where surface (0 to 5 m) nitrate was ≥0.09 µM NO₃ (and throughout the photic-zone <15.00 µM). Prochlorococcus and Synechococcus abundances in the same samples are provided in Supporting Information Table S6. Numbers in parentheses reflect the standard deviation.

© 2017 The Authors. Environmental Microbiology published by Society for Applied Microbiology and John Wiley & Sons Ltd., Environmental Microbiology, 19, 3219–3234
**Bathycoccus** clades have separate niches originated from the comparison of targeted metagenome assemblies to the *B. prasinos* RCC1105 (clade BI) genome (Vaulot et al., 2012; Monier et al., 2013). Follow-up studies have relied on mapping of metagenome and/or metatranscriptome reads to the targeted metagenomes or the RCC1105 genome (Simmons et al., 2016; Vannier et al., 2016). These studies concluded that the two *Bathycoccus* clades rarely co-occur and occupy distinct niches, much like OI and OII. Recent 18S V9 amplicon-based studies (de Vargas et al., 2015; Lopes Dos Santos et al., 2016) cannot discriminate between *Ostreococcus* OI and OII (Monier et al., 2016), or the two *Bathycoccus*, and have yet to be evaluated in the context of abundance data generated by quantitative methods. These amplicon data have been used to infer which green algal taxa dominate in the open ocean, specifically observing that Mamiellophyceae contribute only 1.8% of total photosynthetic sequences and that other prasinophytes, that have fourfold higher relative amplicon percentages, are therefore the dominant group (Lopes Dos Santos et al., 2016). Here, we performed a quantitative enumeration of BI and BII and analysed these in the context of OI and OII abundances and eukaryotic phytoplankton cell counts. Alongside phylogenetic and ultrastructural analyses, our studies establish distributional patterns of these taxa and estimate contributions in different marine regions. The results provide new insights into differentiation and ecological importance of these picophytoplankton that differ from conclusions based purely on omic or amplicon-based relative abundances.

**Phylogenetic relationships and ultrastructural features**

Consistent with a prior study (Monier et al., 2013), analyses of environmental sequences of the ITS1/5.8S/ITS2 (with partial 18S and 28S sequence) documented two statistically supported *Bathycoccus* clades (Fig. 1B). ITS divergence is linked to sexual incompatibility in eukaryotes (Coleman, 2009); hence, the presence of indels in BI and BII ITS sequences and various structural differences (e.g., Supporting Information Fig. S1) could reflect sexual incompatibility (if they prove to be sexual) as a mechanism underlying speciation. This raised the question of whether there are ultrastructural differences between the two *Bathycoccus* clades or the two *Ostreococcus* clades, given that neither BII (e.g., RCC716) nor OII (O. sp. RCC809, Clade B/OII) appear to have been characterized by electron microscopy, and few images exist for *O. lucimarinus* (Clade A/OI) (Worden et al., 2004). The *Ostreococcus* clade OI and OII representatives imaged here (*O. lucimarinus*, *Ostreococcus* sp. RCC809) are more widespread in marine environments than the Clade C (*O. tauri*) and Clade D (*O. mediterraneus*) that have been comprehensively imaged.

Moreover, the QFDEEM method eliminates potential artefacts introduced by standard TEM methodologies used in prior Mamiellophyceae studies (Eikrem and Throndsen, 1990; Chretiennot et al., 1995; Worden et al., 2004; Subirana et al., 2013). Thus, QFDEEM allowed intramembranous particles of the thylakoid interiors to be visualized and revealed details of *Bathycoccus* fine structure, such as nuclear pore morphology, that had not been resolved in published TEM on this organism.

QFDEEM documented strong similarities in the cellular organization and organelle ultrastructure of the four clades studied. Moreover, scale morphology was indistinguishable between BI and BII, as well as a *Bathycoccus prasinos* strain (Eikrem and Throndsen, 1990), and another Mamiellophyceae, *Mantoniella squamata* (Moestrup, 1990). Hence, the environmental preferences of each clade documented here are not obviously reflected by differences in fine structure. One observation of interest relates to scale formation. In larger prasinophytes such as *Pyramimonas* (Moestrup and Walne, 1979) and *Cymbomonas* (Moestrup et al., 2003), scales are readily detected in the cisternae of hypertrophied Golgi apparatus and in secretory vesicles, whereas no scales are evident in the cisternae of the minimal Golgi stacks of *Bathycoccus* (Fig. 2F and G) or *Mantoniella*, nor are secretory vesicles. Instead, we observed embossed scales beneath the *Bathycoccus* membrane and a naked scale just under the cell membrane (Fig. 2I and J), suggesting that the scales may assemble on the inner face of the cell membrane and move through the bilayer to the exterior.

**Different ecotype distributional patterns and environmental influences**

Using our new qPCR assays for quantifying ITS1 sequences from clades BI and BII alongside pre-existing assays for OI and OII, we found that the majority of photic-zone samples contained *Bathycoccus*, *Ostreococcus* or both. Clade BI and clade OI were associated with cooler temperatures, lower salinity and higher nutrient conditions that typically occur in more coastal (shallower) photic-zones (Supporting Information Fig. S7). In contrast, clade BII and clade OII were found in more oligotrophic environments with warmer temperatures, higher salinity, lower nutrients, deeper photic-zones and lower chlorophyll concentrations. Given the extensive geographic overlap and co-occurrence of clade BI and clade OI, we conclude that they both represent ecotypes adapted to mesotrophic environmental conditions.

Overlapping distributions were also observed between clade BI and clade OII. However, there are important biogeographical differences between the clades of each genus. BI and BII were observed at a wider range of conditions and overlapped in a greater number of samples.
(85 samples) than OI and OII (overlapping in 42 samples). This suggests that while OII is adapted to oligotrophic conditions found in oceanic settings (apart from oligotrophic surface waters in strongly stratified water columns), BII tolerates a wider range of environmental conditions. The broader co-existence between Bathycoccus clades is consistent with the fact that they are less evolutionarily divergent than Ostreococcus clades (Simmons et al., 2016), and have greater overlap in their ecological niches.

Regions where Bathycoccus and Ostreococcus ecotypes co-occurred also revealed important factors with respect to variables that might drive distributions of these algae. Clades BII and OII were found at relatively high abundances over a range of depths, from the surface (e.g., Kuroshio Front, and less stratified periods at ALOHA) to 175 m, the maximum depth sampled at ALOHA, suggesting they are not deep-adapted (see, e.g., (Rodriguez et al., 2005)) but instead ‘light-polyvalent’ ecotypes. Depth, temperature and salinity were negatively correlated with abundance of BI and OI, while phosphate and nitrate concentrations, as well as chlorophyll a (OI only) were positively correlated (Supporting Information Table S5). The directionality of these correlations was reversed for BII, except chlorophyll a for which the correlation was not significant, and for OII, where several relationships were not significant. For BII and OII, the k-means clusters provided a more robust sense of environmental conditions where these two taxa appear to thrive (Supporting Information Fig. S7).

Our MBTS study period serendipitously captured a strong warm anomaly that brought record high sea surface temperatures. This anomalous warm patch appeared in the Northeast Pacific at the end of 2013, moved eastward towards the North American coast during May to September 2014 (Bond et al., 2015), and persisted into 2015. It resulted in an expansion of warm subtropical conditions, characterized by low surface concentrations of chlorophyll a, in much of the eastern North Pacific Ocean alongside a deepening of the nutricline and subsurface chlorophyll maximum in the California Current System (Whitney, 2015; Zaba and Rudnick, 2016). These conditions were speculated to have caused reduced primary production. In the MBTS dataset, this anomaly was accompanied by a decrease in nutrient standing stocks and a shift from OI to BI as the dominant clade, and the appearance of OII and BII. Prior studies have not reported OII or BII in Monterey Bay or nearby (Demir-Hilton et al., 2011; Ottesen et al., 2013; Simmons et al., 2016), nor were they present in our Monterey Bay samples from other years. BII reached a maximum (775 ± 71 rDNA copies ml⁻¹) at the end of this period while OII remained very low in abundance (49 ± 40 rDNA copies ml⁻¹). The fact that the appearance of BII and OII coincided with the arrival of the warm patch is consistent with the association of these clades with warm, low-nutrient conditions, and suggests that the anomaly triggered changes in Mamiellophyceae community structure. Prior to the warm anomaly, BI and OI abundances were lowest in Monterey Bay on 5 May when the influence of wind-driven coastal upwelling was strongest, as indicated by cooler sea surface temperatures, and higher nutrient and chlorophyll a concentrations, which favours diatoms in this locale. Low photosynthetic picoeukaryote abundances have been reported off the coasts of Oregon and Chile during similar periods of strong upwelling (Sherr et al., 2005; Collado-Fabbi et al., 2011). Here, as the nutrient-rich water column exhibited more stratification (June), OI bloomed to its maximum while BI abundance remained low (31 ± 6 rDNA copies ml⁻¹; Fig. 3). These results emphasize that each of these picoprasinophyte clades respond to different environmental conditions, or are potentially grazed or lysed by viruses at differential rates.

In contrast to MBTS, Station ALOHA displayed less environmental variability, characteristic of its location in the oligotrophic North Pacific Subtropical Gyre. BII and OII persisted year-round in the DCM and were observed in the surface layer during and just following enhanced wind-mixing (Fig. 4). Their maximum abundances occurred during and after the winter wind-mixing period, in agreement with observations from the Sargasso Sea where timing of prasinophyte blooms appeared to be triggered by 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 was detected by 18S qPCR. In comparison, at Station ALOHA where phosphate concentrations are never as low as those in the Sargasso Sea, BI and OII were observed at relatively high abundances in the DCM throughout the stratified summer months (Fig. 4). As oligotrophic gyres cover vast regions of the planet and support considerable marine primary production (Chavez et al., 2011), the general persistence of Clade II Ostreococcus, and especially Bathycoccus, in gyre habitats suggests these organisms contribute to primary production throughout the year.

These overall distributional and molecular differences support delineation of the Bathycoccus clades BI and BII at the species level, as already determined for the Ostreococcus clades (e.g., Simmons et al., 2016). BI is represented by B. prasinos (Eikrem and Thronsdø, 1990) and was detected here in samples ranging from 8°C to 25°C and salinities of 30 to 35 ppt, but not in our more open-ocean samples. The BII species is represented by the environmental clones sequenced here, targeted metagenomic assemblies published in Monier et al., 2012 and Vannier et al., 2016, ecmarkers defined in Simmons et al., 2016, and understudied isolates RCC715 and RCC716. It is present in warm oligotrophic ocean gyres, with peak abundances typically in well-developed deep-

© 2017 The Authors. Environmental Microbiology published by Society for Applied Microbiology and John Wiley & Sons Ltd., Environmental Microbiology, 19, 3219–3234

Quantitative biogeography of picoprasinophytes 3229
chlorophyll maxima or throughout the photic zone during mixing periods. It overlaps with *B. prasinos*, but extends to warmer, saltier waters ranging from 10°C to 29°C and 33 to 36 ppt among samples analysed here. Finally, nutrient standing stocks were also lower in waters inhabited by BI and OII (Supporting Information Fig. S7), warranting further experimentation on factors that contribute to niche differentiation between the four clades studied here.

**Significant contributions to eukaryotic phytoplankton**

*Bathycoccus* and *Ostreococcus* comprised 34 ± 13% of eukaryotic phytoplankton cells enumerated by flow cytometry at the Station ALOHA DCM, and ranged up to 27% in tropical Atlantic cruise samples (Table 2). These estimates lack contributions from *Micromonas*, which was not enumerated here, and for which there are not yet qPCR assays available that distinguish between its seven clades. This third common Mamiellophyceae genus has been reported at up to 700 cells ml\(^{-1}\) in the open-ocean (Treusch et al., 2012) and higher abundances in coastal environments (10 000 cells ml\(^{-1}\), see for example; Collado-Fabbri et al., 2011). Assuming the genome-sequenced representatives of the Mamiellophyceae accurately reflect the rRNA-transcriptional-unit copy numbers (i.e., two copies) across the members of each clade analysed herein, our results contrast with conclusions drawn from relative abundances derived from 18S V9 rRNA amplicon data, particularly for the DCM. Mamiellophyceae formed only 1.8% of the total photosynthetic sequences in Tara Oceans data (Lopes Dos Santos et al., 2016). In Tara pico/nano-planktonic size-fractionated (<5 μm) DCM samples, Mamiellophyceae comprised 7% of photosynthetic eukaryote sequences and 11% in Ocean Sampling Day surface data, which comes largely from coastal sites where Mamiellophyceae are often highly abundant based on FISH and qPCR data (Worden and Not, 2008; Collado-Fabbri et al., 2011; Massana, 2011; Simmons et al., 2016).

Interpretation of amplicon-based relative abundances is complicated by large variations in 18S rRNA gene copies among eukaryotes (Prokopowich et al., 2003; Godhe et al., 2008) as well as potential primer biases. While genome-sequenced representatives of the Mamiellophyceae analysed here have two rRNA gene (and ITS) copies, copy numbers in larger eukaryotes vary by orders of magnitude and a documented correlation exists between eukaryotic phytoplankton cell size and copy number (Zhu et al., 2005; Godhe et al., 2008). Thus, the extent to which amplicon-based percent contributions reflects organismal abundances is often unclear and this issue is exacerbated when comparing across different phyla. In the samples investigated here, seasonal and site-to-site variations in both eukaryotic cell counts and those of *Bathycoccus* and *Ostreococcus* were large (with low technical variation for each measurement), making the consistent, temporally-resolved sampling at the time-series sites exceptionally valuable for determining trends and taxon contributions. Overall, *Bathycoccus* and *Ostreococcus* together averaged around 20% of the eukaryotic phytoplankton, except in open-ocean surface waters during stratified periods when contributions were much lower (Table 2).

**Conclusions**

We have demonstrated distinct biogeographies for *Bathycoccus* and *Ostreococcus* ecotypes that appear to partition at the species level. We show that the OI and BI ecotypes, represented by *O. lucimarinus* and *B. prasinos*, respectively, are present in cooler, mesotrophic ecosystems. In contrast, the BI ecotype, which awaits formal description as a species, is present in warm oligotrophic waters, but has an expanded temperature range relative to OII (represented by *O. sp. RCC809*) because it extends into cooler, more nutrient-rich waters. From an oceanographic perspective, the large seasonal variations in organismal abundances observed here emphasize the importance of time-series sampling for understanding phytoplankton distributions. Importantly, these two Mamiellophyceae genera contributed significantly to total eukaryotic phytoplankton counts in the marine ecosystems studied, including the open ocean. Thus, while amplicon sequencing is powerful for investigating diversity and distributions, this study reinforces the importance of using absolute quantification methods to establish organismal abundance patterns. Our findings suggest that the long-term trajectories of the different *Bathycoccus* clades studied here, and the two *Ostreococcus*, will differ as surface temperatures and stratification patterns shift under future ocean conditions.

**Experimental procedures**

**Sample collection**

Samples were collected from temperate to tropical regions on multiple cruises (Table 1). The entire dataset had 259 samples, including 78 from station ALOHA in the North Pacific Subtropical Gyre and 22 from the Monterey Bay Time-series (MBTS) station M1. Niskin bottles affixed to a Conductivity, Temperature and Depth (CTD) rosette sampler were used to collect seawater (Sea-Bird Electronics, Bellevue, WA, USA). NO\(_3\), NO\(_2\), PO\(_4\)\(^3-\), Si(OH)\(_4\) of which generally 500 to 2000 mL were filtered for DNA samples (details are provided in the Supporting Information). Chlorophyll a and nutrient samples were collected and analysed according to previously described procedures (Pennington and Chavez, 2000; Santoro et al., 2013). For 190 of the 259 DNA samples, flow cytometry samples were preserved and analysed on an InFlux as described in (Cuvelier et al., 2010). These samples were run at ~25 ml min\(^{-1}\) and weighed to establish the volume run. Additional
Station ALOHA metadata and methods are available at http://hahana.soest.hawaii.edu/hot/hot-dogs/.

**Culturing, estimating flow cytometry error and electron microscopy**

*Bathycoccus prasinus* CCMP1898 and *O. lucimarinus* CCMP2972 were acquired from National Center for Marine Algae and Microbiota (East Boothbay, ME, USA). *Bathycoccus* RCC715, RCC716 and *O. sp. RCC809 were acquired from the Roscoff Culture Collection (Roscoff, France) and grown in semi-continuous batch cultures in L1 media (Guillard, from the Roscoff Culture Collection (Roscoff, France) and size Supor filters and flash frozen in liquid N2. For transmis-

Flow cytometry error rates were evaluated by diluting cul-
tures of RCC716 and CCMP2972 to environmentally-relevant concentrations, fixing them as done for field samples and running them on the InFlux (as for field samples) on three independent days. The percent variation for cell counts ml^{-1} was 1.18% (*Bathycoccus*) and 1.65% (*Ostreococcus*). We also analysed a field sample on three independent days, render-
ing variations of 1.71% (*Prochlorococcus*), 0.96% (*Synechococcus*) and 2.86% (eukaryotic phytoplankton).

**DNA extraction, clone libraries and quantitative real-time polymerase chain reaction**

Methods for DNA extraction and clone library construction are provided in Supporting Information as are additional details of BI and BII primer-probe design and testing (Supporting Information Tables S2 and S3). Briefly, for clone libraries, a general eukaryote reverse primer that anneals to a conserved region at the beginning of the 28S rRNA gene (ITS055R: 5'-CTCCTTGGTCCGTTTCAGACGCGG-3') (Marin et al., 1998) was used alongside a forward primer (MAM18S1F: 5'CGAAAGGCTCTRGGATGTCCTTCG-3') and flanking partial 18S rRNA (148 bp), ITS1, 5.8S, ITS2 and flanking partial 28S (14 bp) were aligned using MAFFT (Katoh and Standley, 2013) with default settings. Identical sequences from the same clone library were removed and a maximum likelihood reconstruction was performed using 664 nucleotide positions (10% gap tolerance) in MEGA6 (Tamura et al., 2013). A general time reversible model with invariant sites and a gamma rate distribution (GTR + I + G) was used. Bootstrap percentages were calculated in RaxML (Stamatakis, 2014) and PhyML 3.1 (Guindon et al., 2010) from 1000 replicates.

**Phylogenetic analysis**

Phylogenetic analyses utilized representative clone library sequences generated here and sequences from targeted metagenomes and GenBank. Forty-seven sequences covering the flanking partial 18S rRNA (148 bp), ITS1, 5.8S, ITS2 and flanking partial 28S (14 bp) were aligned using MAFFT and 1000 replicates. A general time reversible model with invariant sites and a gamma rate distribution (GTR + I + G) was used. Bootstrap percentages were calculated in RaxML (Stamatakis, 2014) and PhyML 3.1 (Guindon et al., 2010) from 1000 replicates.

**Statistical analyses of qPCR and environmental data**

As qPCR and environmental data failed tests for normality, non-parametric statistical tests were chosen to assess significant relationships. Matrices contained qPCR data, temperature, salinity, density (ρ), depth, NO\textsubscript{3}, NO\textsubscript{2}, NH\textsubscript{4}, PO\textsubscript{4}, and chlorophyll a, and were assembled with all samples as well as for individual transects and time-series studies to evaluate distinct regimes. Non-metric multidimensional Scaling (NMDS) analysis was performed on a dissimilarity matrix containing 213 samples and 6 variables (temperature, salinity, depth, NO\textsubscript{3}, PO\textsubscript{4}, N:P ratio). Data were z-scored to standardize units before using Euclidean distances to calculate the dissimilarity matrix. NMDS coordinates were used in a k-means cluster analysis (based on square Euclidean distance, 50 replications) to divide samples into 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). Temperature and salinity distributions of *Bathycoccus* and *Ostreococcus* clades included samples where a clade was detected but not quantifiable (i.e., <10 rDNA copies well^{-1}).

**Acknowledgements**

We thank the captains and crews of the R/Vs Western Flyer, Rachel Carson, Melville, Thompson, Seward Johnson, Nat-ushima and Kilo Moana. We also thank Francisco Chavez and Timothy Pennington for cruise support, Adrian Marchetti.
References

Bond, N.A., Cronin, M.F., Freeland, H., and Mantua, N. (2015) Causes and impacts of the 2014 warm anomaly in the NE Pacific. Geophys Res Lett 42: 3414–3420.

Chavez, F.P., Messie, M., and Pennington, J.T. (2011) Marine primary production in relation to climate variability and change. Annu Rev Mar Sci 3: 227–260.

Chretiennot, M.J., Courties, C., Vaquer, A., Neveux, J., Claustre, H., Lautier, J., and Machado, M.C. (1995) A new marine picoeukaryote: Ostreococcus tauri gen. et sp. nov. (Chlorophyta, Prasinophyceae). Phycologia 34: 285–292.

Clayton, S., Lin, Y.-C., Follows, M.J., and Worden, A.Z. (2017) Co-existence of distinct Ostreococcus ecotypes at an oceanic front. Limnol Oceanogr 62: 75–88.

Coleman, A.W. (2009) Is there a molecular key to the level of “biological species” in eukaryotes. Mol Phylogeny Evol 50: 197–203.

Collado-Fabbri, S., Vaulot, D., and 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, P.D., and Caron, D.A. (2006) Abundance and distribution of Ostreococcus sp. in the San Pedro Channel, California, as revealed by quantitative PCR. Appl Environ Microbiol 72: 2496–2506.

Cuvelier, M.L., Allen, A.E., Monier, A., McCrow, J.P., Messié, M., Tringe, S.G., et al. (2010) Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. Proc Natl Acad Sci USA 107: 14678–14683.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., et al. (2015) Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. Science 348: 126105.

Demir-Hilton, E., Sudek, S., Cuvelier, M.L., Gentemann, C., Zehr, J.P., and Worden, A.Z. (2011) Global distribution patterns of distinct clades of the photosynthetic picoeukaryote Ostreococcus. ISME J 5: 1095–1107.

Eikrem, W., and Thronsden, J. (1990) The ultrastructure of Bathycoccus gen. nov. and Bathycoccus prasinus sp. nov., a non-motile picoplanktonic alga (Chlorophyta, Prasinophyceae) from the Mediterranean and Atlantic. Phycologia 29: 344–350.

Godhe, A., Asplund, M.E., Harnstrom, K., Saravanan, V., Tyagi, A., and Karunasaraj, I. (2008) Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. Appl Environ Microbiol 74: 7174–7182.

Gomez-Pereira, P.R., Hartmann, M., Grob, C., Tarran, G.A., Martin, A.P., Fuchs, B.M., et al. (2013) Comparable light stimulation of organic nutrient uptake by SAR11 and Prochlorococcus in the North Atlantic subtropical gyre. ISME J 7: 603–614.

Goodenough, U.W., and Staehelin, L.A. (1971) Structural differentiation of stacked and unstacked chloroplast membranes. Freeze-etch electron microscopy of wild-type and mutant strains of Chlamydomonas. J Cell Biol 48: 594–619.

Guillard, R.R.L. (1975) Culture of phytoplankton for feeding marine invertebrates. In Culture of Marine Invertebrate Animals. Smith, W.L., and Chanley, M.H. (eds). New York, USA: Plenum, pp. 29–60.

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.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol 59: 307–321.

Henderson, G.P., Gan, L., and Jensen, G.J. (2007) 3-D ultrastructure of O. tauri: Electron cryotomography of an entire eukaryotic cell. PLoS One 2: e749.

Hessler, J.E. (2011) The origins and evolution of freeze-etch electron microscopy. J Electron Microsc (Tokyo) 60: S3–29.

Katoh, K., and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30: 772–780.

Lopes Dos Santos, A., Gourvil, P., Tragin, M., Noel, M.H., Decelle, J., Romac, S., and Vaulot, D. (2016) Diversity and oceanic distribution of prasinophytes clade VII, the dominant group of green algae in oceanic waters. ISME J 11: 512–528.

Marin, B., and 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.

Marin, B., Klingberg, M., and 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.

Massana, R. (2011) Eukaryotic picoplankton in surface oceans. Annu Rev Microbiol 65: 91–110.

Metfies, K., von Appen, W.J., Killas, E., Nicolaus, A., and Nothig, E.M. (2016) Biogeography and Photosynthetic Biomass of Arctic Marine Pico-Eukaryotes during Summer of the Record Sea Ice Minimum 2012. PLoS One 11: e0148512.

Moestrup, O. (1990) Scale structure in Mantonielia squamata, with some comments on the phylogeny of the Prasinophyceae (Chlorophyta). Phycologia 29: 437–442.

Moestrup, O., and Walne, P.L. (1979) Studies on scale morphogenesis in the Golgi apparatus of Pyramimonas tetrarhynchus (Prasinophyceae). J Cell Sci 36: 437–459.

Moestrup, O., Inouye, I., and Hori, T. (2003) Ultrastructural studies on Cymbomonas tetrarhynchus (Prasinophyceae). I. General structure, scale microstructure, and ontogeny. Can J Bot 81: 657–671.

Monier, A., Welsh, R.M., Gentemann, C., Weinstock, G., Sodergren, E., Armbrust, E.V., et al. (2012) Phosphate...
transporters in marine phytoplankton and their viruses: Cross-domain commonalities in viral-host gene exchanges. *Envir Microbiol* **14**: 162–176.

Monier, A., Sudek, S., Fast, N.M., and Worden, A.Z. (2013) Gene invasion in distant euarktyotic lineages: Discovery of mutually exclusive genetic elements reveals marine biodiversity. *ISME J* 7: 1764–1774.

Monier, A., Findlay, H.S., Charvet, S., and Lovejoy, C. (2014) Late winter under ice pelagic microbial communities in the high Arctic Ocean and the impact of short-term exposure to elevated CO2 levels. *Front Microbio* 5: 490.

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

Moreau, H., Verhelst, B., Couloux, A., Derelle, E., Rombouts, 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., 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., Balague, V., et al. (2008) Protistan assemblages across the Indian Ocean, with a specific emphasis on the picoeukaryotes. *Deep-Sea Res Part I Oceanogr Res Pap* **55**: 1456–1473.

Ottesen, E.A., Young, C.R., Eppley, J.M., Ryan, J.P., Chavez, F.P., Scholin, C.A., and DeLong, E.F. (2013) Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc Natl Acad Sci USA* **110**: E488–E497.

Pennington, J.T., and Chavez, F.P. (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 Part II* **47**: 947–974.

Prokopowich, C.D., Gregory, T.R., and Crease, T.J. (2003) The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **46**: 48–50.

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

Santoro, A.E., Sakamoto, C.M., Smith, J.M., Plant, J.N., Gehman, A.L., Worden, A.Z., et al. (2013) Measurements of nitrite production and nitrite-producing organisms in and around the primary nitrite maximum in the central California Current. *Biogeosci Discuss* **20**: 5803–5840.

Sherr, E.B., Sherr, B.F., and Wheeler, P.A. (2005) Distribution of cocoid cyanobacteria and small eukaryotic phytoplankton in the upwelling ecosystem off the Oregon coast during 2001 and 2002. *Deep-Sea Res Part II* **52**: 317–330.

Simmons, M.P., Sudek, S., Monier, A., Limardo, A.J., Jimenez, V., Perle, C.R., et al. (2016) Abundance and biogeography of picoprasinophyte ecotypes and other phytoplankton in the Eastern North Pacific Ocean. *Appl Environ Microbiol* **82**: 1693–1705.

Stamatakis, A. (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Subirana, L., Pequin, B., Michely, S., Escande, M.L., 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., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**: 2725–2729.

Treu, A.H., Demir-Hilton, E., Vergin, K.L., Worden, A.Z., Carlson, C.A., Donatz, M.G., et al. (2012) Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. *ISME J* **6**: 481–492.

van Baren, M.J., Bachy, C., Reistetter, E.N., Purvine, S.O., Grimwood, J., Sudek, S., et al. (2016) Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC Genomics* **17**: 267.

Vannier, T., Leconte, J., Seeleuthner, Y., Mondy, S., Pelletier, E., Aury, J.M., et al. (2016) Survey of the green picoalga *Bathycoccus* genomes in the global ocean. *Sci Rep* **6**: 37900.

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, F.A. (2015) Anomalous winter winds decrease 2014 transition zone productivity in the NE Pacific. *Geophys Res Lett* **42**: 428–431.

Worden, A.Z. (2006) Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquat Microbial Ecol* **43**: 165–175.

Worden, A.Z., and Not, F. (2008) Ecology and diversity of picoeukaryotes. In *Microbial Ecology of the Oceans*. Kirchman, D.L. (ed). Hoboken: Wiley, p. 594.

Worden, A.Z., Nolan, J.K., and Palenik, B. (2004) Assessing the dynamics and ecology of marine picophytoplankton: The importance of the eukaryotic component. *Limonol Oceanogr* **49**: 168–179.

Zaba, K.D., and Rudnick, D.L. (2016) The 2014–2015 warming anomaly in the Southern California Current System observed by underwater gliders. *Geophys Res Lett* **43**: 1241–1248.

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

**Supporting information**

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

**Fig. S1.** Secondary structure of the BI and BII Internal Transcribed Spacers (ITS2) in *Bathycoccus prasinos* CCMP1898 (JX625115) representing BI and environmental clone TAM24 (KY368638) as a representative sequence for clad BI. Note that within clad BI identity is 94%-100% and 97%-100% for the ITS1 and ITS2, and 96%-100% and 98%-100% for BI, respectively. The between clad identities are 97%-100% and 96%-100% and 98%-100% for BI and BII, respectively. * indicates a potential SNP (or PCR/sequencing error) and v symbols indicate a potential single nucleotide insertion/deletion. Color

© 2017 The Authors. Environmental Microbiology published by Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*, **19**, 3219–3234
coding reflects three types of nucleotide changes present in multiple clones within clade BI, the same was not observed for BI clones. (c) In addition to differences noted in the main text for BI (based on TAM24; also present in RCC716, KY563784) the ITS2 universal Helix 1 of BI sequences (represented by B. prasinos CCMP1898) has a 10 nt loop composed of 5'-CUUUUAUUUUU-3' from positions 2104–2113 in JX625115. (d) Close up comparison of the Bathycocccaceae family-specific additional ITS2 helix, located between universal Helices 3 and 4 (Marin and Melkonian, 2010). Also note (not shown) that base pairs 13 and 19 of universal Helix 2 are G-C and C-G (base pairs 678–703 and 684–697, respectively in KY368638), instead of A-C and U-G (base pairs 2138–2159 and 2144–2153, respectively in JX625115).

Fig. S2. Alignment of Mamiellophyceae ITS sequences. QPCR primer regions targeted by assays designed in this study for Bathycoccus clades BI and BI are highlighted by black bold lettering, and probe sequences are in blue. Sequences shown originate from B. prasinos CCMP1898 (clade BI, JX625115), RCC716 (Bathycoccus clade BI, KY563784), CCE9901 (O. lucimarinus, clade OI, extracted from genome reference sequence CP000588), RCC809 (O. sp., clade OII, extracted from JGI genome v2.0), OTH95 (O. tauri, extracted from JGI genome v2.0), RCC2583 (O. mediterraneus, JN862916), RCC299 (M. commoda, extracted from genome reference sequence CP001575) and CCM1545 (M. pusilla, KY954994). Note that BBAN7 (used in several prior publications) and RCC1105 are synonymous.

Fig. S3. Ostreococcus (a) clade OI and (b) OII. The etched true surface of the cell membrane is smooth (S) in OI and carries knobs (K) in OII. (c) Bathycoccus clade BI. The etched true surface of the cell membrane, populated by four scales (asterisks), is smooth (s), similar to the surface of clade OI (a). (d) Bathycoccus clade BI, with layered scales exposed by etching. (e) Bathycoccus clade BI showing scale cross-sectional thickness. (f) Bathycoccus clade BI cells visualized by glutaraldehyde/OsO₄ fixation, dehydration, plastic embedding, thin-sectioning and lead and uranyl staining. The uneven layering of the scales is evident. (g) Bathycoccus clade BI, thin section. Arrow points to a scale indistinguishable from the negatively-stained BI scale shown in Fig. 2K. Asterisks mark scales that adopt a slightly cupped shape. (h) Bathycoccus clade BI. Asterisks mark scales that adopt a slightly cupped shape. Scale bars: 100 nm.

Fig. S4. Results of Bathycoccus qPCR assays using mixtures of genomic DNA from B. prasinos CCMP1898 (clade BI) and Bathycoccus RCC716 (clade BI).

Fig. S5. Abundance 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 BI clades were detected (x) are indicated for both WFAD09 and GOC12. (b) Abundances of Bathycoccus and Ostreococcus clades by qPCR at different depths at each station during the GOC12 transect cruise. Note the logarithmic scaling for abundances (heat map).

Fig. S6. Clade ratios (Bathycoccus BI:BI or Ostreococcus OI:OII) of abundances determined by qPCR at stations where ecotypes co-occurred on WFAD09 and GOC12 cruises. Ratios are plotted against (a) depth and (b) PO₄³⁻ on y-axis. Ratios are plotted against depth on y-axis. Note that ecotype ratios are shown on a logarithmic scale and only data from samples where both Ostreococcus clades or both Bathycoccus clades were detected are included.

Fig. S7. (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 quartile and above third quartile are also shown (circles). (c) Table of mean and standard deviation for various environmental parameters associated with each clade. Clusters 2 and 3 contained higher abundances of clades BI and OI than Cluster 1. 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 for samples where at least one clade was detected 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. Abundance of Bathycoccus and Ostreococcus clades across the three environmental clusters were also compared by Kruskal-Wallis tests with Dunn’s post hoc tests. No clade showed a significant difference in abundance between Clusters 2 and 3 using this statistical test. However, BI and OI abundances were significantly (p<0.01) lower in Cluster 1 than in Clusters 2 and 3. Moreover, clades BI and OII were significantly higher in Cluster 1 compared to the other clades.

Table S1. Results from environmental clone libraries targeting Mamiellophyceae and cultured strains.

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

Table S3. 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. AFD, cultured isolates of the BI clade that awaits formal description as a new species.

Table S4. Mean and standard deviation of environmental parameters at Station ALOHA in samples where both Bathycoccus BI and Ostreococcus OII were detected, and where neither genus was observed.

Table S5. Spearman’s rank correlation coefficient (p) between ecotype qPCR abundances and environmental parameters. * indicates correlation not significant (p > 0.01).

Table S6. Abundance of Prochlorococcus and Synechococcus in the same regions as for Table 2. *Note that 25 m above the DCM at Station ALOHA Prochlorococcus was always more than twice as abundant (150 856 ± 30 644 annual average) than at the DCM. Latitudes for samples used in these averages are between 6°N and 37°N.

Supplementary Database
Table compiling qPCR results and metadata of the 259 samples examined in this study.