

Portland State University

PDXScholar

Biology Faculty Publications and Presentations

Biology

1-28-2021

Niche Partitioning of Low-Light Adapted Prochlorococcus Subecotypes Across Oceanographic Gradients of the North Pacific Subtropical Front

Anne W. Thompson

Portland State University, anne.w.thompson@pdx.edu

Kathleen Kouba

Oregon State University

Nathan A. Ahlgren

Clark University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/bio_fac



Part of the [Biology Commons](#), and the [Oceanography Commons](#)

Let us know how access to this document benefits you.

Citation Details

Thompson, A. W., Kouba, K., & Ahlgren, N. A. (2021). Niche partitioning of low-light adapted *Prochlorococcus* subecotypes across oceanographic gradients of the North Pacific Subtropical Front. *Limnology and Oceanography*, Ino.11703. <https://doi.org/10.1002/Ino.11703>

This Article is brought to you for free and open access. It has been accepted for inclusion in Biology Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Niche partitioning of low-light adapted *Prochlorococcus* subecotypes across oceanographic gradients of the North Pacific Subtropical Front

Anne W. Thompson¹,^{*} Kathleen Kouba,² Nathan A. Ahlgren³

¹Department of Biology, Portland State University, Portland, Oregon

²College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon

³Biology Department, Clark University, Worcester, Massachusetts

Abstract

The cyanobacterium *Prochlorococcus* is the most abundant photosynthetic cell on Earth and contributes to global ocean carbon cycling and food webs. *Prochlorococcus* is known for its extensive diversity that falls into two groups of ecotypes, the low-light (LL) and high-light (HL) adapted ecotypes. Previous work has shown niche partitioning of the very abundant HL adapted ecotypes and subecotypes across oceanographic gradients including temperature, nutrients, and day length. However, niche partitioning of subecotypes within the LL adapted ecotypes has not been studied as well because they are less abundant and less accessible than surface, HL populations. Through high-throughput, cyanobacterial-specific, sequencing of high-resolution depth profiles from three stations across the North Pacific Subtropical Front, we discovered extensive diversity and strong latitudinal partitioning of subecotype populations within LL ecotypes, in contrast to fairly consistent patterns at the broader level of LL *Prochlorococcus* ecotypes. These results indicate that both shallow and deep-water processes shape microbial community structure in this region despite strong connectivity in microbial communities across the North Pacific Subtropical Front, supported by a few HL and LL sequences found ubiquitously across stations. This study expands understanding of the diversity and complexity of *Prochlorococcus* community structure over oceanographic gradients.

The open-ocean picocyanobacterium *Prochlorococcus* is the most abundant photosynthetic organism on Earth and impacts biogeochemical cycles on global scales (Chisholm et al. 1992; Partensky et al. 1999; Biller et al. 2015). The stability of large *Prochlorococcus* populations is remarkable over vast oceanographic gradients (Johnson et al. 2006; Flombaum et al. 2013; Ribalet et al. 2015), ephemeral mixing events (Thompson et al. 2018), and seasons (Malmstrom et al. 2010).

The extensive diversity of the *Prochlorococcus* “collective” (the entire diverse lineage) is one explanation for the stability of these cell populations across strong gradients and dynamic

oceanic conditions (Biller et al. 2015; Kashtan et al. 2017). Differences in light, nutrient, and mortality among coexisting cells may mean subsets of cells are genomically equipped to thrive in new conditions or over large gradients. Thus, the population as a whole maintains abundance and function across a range of conditions, gradients, and geographical regions.

At coarse taxonomic resolution, *Prochlorococcus* is composed of 12 known clades (or ecotypes). The high-light (HL) adapted ecotypes, named for their preference to higher light intensities (Moore et al. 1998), share affinity for high light environments but are differentiated by temperature preferences (Johnson et al. 2006; Chandler et al. 2016). The most globally abundant ecotypes, including some of the “low light” (LL) adapted ecotypes (HLI, HLII, LLI, LLII/III, LLIV), show distinct and complementary distributions over depth (Zinser et al. 2006, 2007; Malmstrom et al. 2010), seasonal cycles (Malmstrom et al. 2010), nutrients (Larkin et al. 2016, 2019), and iron (Rusch et al. 2010).

Extensive *Prochlorococcus* diversity also exists at finer taxonomic scales, below the ecotype level. Several studies have shown that HL subecotypes are controlled by unique combinations of environmental variables including temperature, day length, and phosphate (Kent et al. 2016; Larkin et al. 2016; Ahlgren et al. 2019). *Prochlorococcus* HLII subecotype patterns are also apparent at the single cell level. In

*Correspondence: anne.w.thompson@gmail.com

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Additional Supporting Information may be found in the online version of this article.

Author Contribution Statement: A.W.T. conceived of the research project and carried out data analysis and writing. K.K. carried out sampling during the cruise, chose sampling locations based on oceanographic conditions, processed samples, and contributed to the writing of the manuscript. N.A.A. participated in data analysis, interpretation, statistical analysis, and writing of the manuscript.

terms of gene content, single cells form cohesive subpopulations that have distinct and complementary shifts in abundance with seasonal changes in ocean conditions (Kashtan et al. 2014, 2017). Similarly, subecotypes of the closely related picocyanobacterium *Synechococcus* have distinct abundance patterns across oceanographic gradients and time (Tai et al. 2011; Farrant et al. 2016; Ahlgren et al. 2019).

Despite the growing number of studies that look at HL ecotype microdiversity, the ecology and environmental drivers of LL ecotypes and LL subecotypes remain poorly studied due to their low abundance in easily accessed surface waters. Representative LL *Prochlorococcus* are more genetically varied than HL *Prochlorococcus* (Rocap et al. 2002; Ahlgren and Rocap 2006; Lavin et al. 2008). With larger genomes (Rocap et al. 2003; Kettler et al. 2007; Biller et al. 2014), LL *Prochlorococcus* have more genome-encoded options for nutrient acquisition (Martiny et al. 2009a), more complex light physiology (Bibby et al. 2003), and possibly different sensitivities to mortality sources such as through viral lysis (Sullivan et al. 2003), interactions with heterotrophic bacteria (Sher et al. 2011), and production of lantipeptides (Li et al. 2010). Some studies suggest that LL *Prochlorococcus* should be classified as distinct species (Moldovan and Gelfand 2018) or genera (Tschoeke et al. 2020) apart from the HL clades. Measuring community structure within LL ecotypes, and improving understanding of the environmental factors that drive the ecology of LL ecotypes and subecotypes, will improve understanding of *Prochlorococcus* ecology, taxonomy, and contributions to biogeochemical cycling.

In this work, we examine the ecology and dynamics of *Prochlorococcus* subecotypes, especially within LL-clades, by measuring 16S-23S rRNA internal transcribed spacer (ITS) diversity over 12 high-resolution depth profiles (5 m depth intervals from surface to 145 m) at three stations spanning the North Pacific Subtropical Front. Our study region includes sampling within the North Pacific Subtropical Gyre, at the frontal boundary, and within the North Pacific Transition Zone, where dramatic gradients in picocyanobacterial community structure have been observed in surface waters (Ribalet et al. 2015; Gradoville et al. 2020). We found high diversity among LL clades, especially within the enigmatic LLI clade. Like the HL clades, LL ecotypes contained subecotypes with distinct environmental drivers. This work expands what is known about the role of diversity in the *Prochlorococcus* and supports the hypothesis that extensive diversity enables the stability of large picocyanobacterial populations over oceanographic gradients.

Methods and materials

Oceanographic sampling

Samples were collected at three locations (named by their latitudes) on either side and within the North Pacific Subtropical Front in March 2017 as part of the Submesoscale Mixed Layer Experiment (SMILE) designed to improve understanding

of submesoscale (1–10 km) physical processes in the mixed layer (cruise SKQ201703S). This region was chosen because it is dominated by submesoscale processes with relatively weak mesoscale eddies. This region is typical of large areas of the Earth's open oceans. Station locations are as follows: "26N" (26°N; 147.1°W), "30N" (30.5°N; 145.6°W), and "35N" (35.1°N; 139.58°W). For depth profiles, a 24-bottle CTD rosette was used to collect seawater water samples at 5 m intervals between the surface (5 m) and the base of the euphotic zone (145 m). Conductivity, temperature, salinity, and oxygen were measured by instruments mounted on the CTD rosette, calibrated at the beginning of the cruise. In addition, sea surface temperature data from the Moderate Resolution Imaging Spectroradiometer (MODIS) were used for larger-scale oceanographic context. NASA SeaDAS was used to visualize satellite data. The *oce* package in R was used to process oceanographic data from CTD files (Kelley 2018).

Nutrients

Samples for nutrients were collected at approximately 20–25 m intervals over all depth profiles. Nutrients analyzed included the sum of nitrite (NO_2^-) plus nitrate (NO_3^-) (N + N), ammonium (NH_4), phosphate (PO_4), and silicate (Si). Nutrient samples were processed at the Oregon State University Elemental Analyzer Facility with a hybrid instrument. Analysis was carried out with a continuous flow analysis system including two channels of an Alpkem AutoAnalyzer II (configured with a 5 cm optical path), used for PO_4 and NH_4 ; and three channels of an Astoria Pacific Rapid Flow Analyzer for Si, NO_3^- , and NO_2^- .

Flow cytometry analysis

Duplicate samples for flow cytometry were collected within 1 h of collecting water from the CTD by fixation of 2 mL seawater with a final concentration of 0.125% glutaraldehyde (Tousimis, Rockville, Maryland), incubation for 10 min at room temperature, then flash freezing in liquid nitrogen and storage at -80°C until processing. Chlorophyll-containing cells (i.e., phytoplankton) were quantified by flow cytometry. Cells from thawed samples were interrogated with a 488 nm laser using a BD Influx flow cytometer (BD Biosciences, San Jose, California). Groups of phytoplankton were distinguished based on their relative red fluorescence (chlorophyll, bandpass 692/40 nm), relative orange fluorescence (phycoerythrin, bandpass 572/27 nm), side scatter, and forward scatter signals. The trigger for data collection was forward scatter. *Synechococcus* populations were identified as cells with both chlorophyll and phycoerythrin fluorescence. *Prochlorococcus* cells were identified as cells with relatively low forward scatter signals, relatively low chlorophyll fluorescence signals, and no phycoerythrin fluorescence. Pigmented picoeukaryotes (PPEs) were identified as cells with relatively high chlorophyll fluorescence, high forward scatter, and no phycoerythrin. Gating was carried out using FlowJo. Counts of cells measured by flow

cytometry were normalized to the volume of seawater analyzed. Flow cytometry data has been archived with the Biological and Chemical Oceanography Data Management Office (Thompson 2020).

Classifying amplicon sequence variants

Replicates of 100 mL of seawater from each cast and depth were filtered onto 0.2 μm polycarbonate filters within 1 h of collection from the CTD. Filters were archived in 2 mL bead-beater tubes, with equal amounts of 0.1 and 0.5 mm glass beads, until return to the lab for sample processing. DNA was extracted using Qiagen's DNeasy kit (Qiagen, Hilden, Germany) with the addition of 2 min of bead-beating at 30 Hz prior the manufacturer's protocol. Cyanobacterial community composition was assessed by amplicon sequencing of the 16S-23S rRNA ITS region, which provides subecotype taxonomic resolution. Primers were linked to Illumina adapter sequences and unique "indexes," as described in (Ahlgren et al. 2019). Triplicate polymerase chain reactions (PCRs) with 2 ng template DNA and the uniquely indexed primer pairs were performed for each DNA sample. Thermocycling parameters included initiation and denaturation at 95°C for 2 min, then 25 cycles of 95°C for 45 s, 55°C for 45 s, and 68°C for 90 s, followed by an elongation step of 68°C for 10 min using Platinum Taq DNA Polymerase High Fidelity (ThermoFisher Scientific) with primers at a final concentration of 400 nmol L⁻¹. Triplicate PCRs were pooled, quantified, then purified with Axygen™ AxyPrep Mag™ PCR Clean-up Kits (Axygen Scientific, Union City, California, U.S.A.), quantified by Qubit Fluorometric Quantification (Qubit, ThermoFisher Scientific), and quantified by quantitative polymerase chain reaction (qPCR) targeting Illumina adaptors (Kapa Biosystems, Roche). Equal amounts of the PCR product from each template were pooled to create the multiplexed sample library for sequencing. Sequencing was carried out by Illumina MiSeq (2 × 300 bp paired end reads). Sequence reads were demultiplexed using QIIME1 (Caporaso et al. 2010). Demultiplexed reads were processed by *dada2* (Callahan et al. 2016) to remove low quality reads, model error relative to quality, remove chimeras, and identify unique amplicon sequence variants (ASVs). Due to short sequencing reads relative to the target region, forward and reverse reads could not be merged, and only reverse reads were used in final analysis as they cover the ITS-1 region and provide more discrimination between cyanobacteria. Sequence reads were truncated at the first instance of a quality score less than 11 (*truncQ*) and *maxEE* (expected errors) was set to 1. A total of 11,914,931 quality-filtered sequence reads were used to determine ASVs. The ecotype-level identity of each unique ASV with more than 100 sequence reads across the entire data set was identified using BLASTn (Altschul et al. 1990) against a database of marine cyanobacterial sequences of known ecotype identity ($n = 4230$ ASVs). Ecotype identity of references was determined from previous phylogenetic analyses of 16S-23S ITS regions (Rocap et al. 2002; Lavin

et al. 2008; Huang et al. 2011; Biller et al. 2014) and phylogenetic analyses of single-cell amplified genomes (Berube et al. 2018). ASVs with a best BLAST hit to a *Prochlorococcus* sequence, but less than 90% sequence identity, were categorized *Prochlorococcus* "low light unclassified" or "high light unclassified" depending on the ecotype identity of their best BLASTn hit. The 90% similarity threshold was chosen based on similarities between phylogenetically well-characterized ecotype ITS regions (Rocap et al. 2002). Ecotype-level community structure was determined from relative read abundance of ASVs grouped by BLASTn ecotype identity in an approach validated previously by qPCR-determined gene copy counts (Larkin et al. 2016). Sequence data are available via GenBank (#PRJNA636778).

Rarefaction

Tables of counts per ASV were processed in R and R Studio. Rarefaction was performed using the *vegan* package (Oksanen et al. 2019) function *rrarefy*. The effect of rarefaction on alpha (*diversity* with *vegan*) and beta (*vegdist* with *vegan*) diversity was tested for different sample sizes (i.e., levels of rarefaction) ranging between 500 and 20,000 sequences per sample. For alpha diversity, the difference between five replicates of the unrarefied (whole) data set and five replicates of each rarefied data set was tested with the Student's *t*-test (variance estimated separately for both groups and the Welch modification to the degrees of freedom). For beta diversity, the difference between five replicates of the unrarefied (whole) data set and five replicates of each rarefied data set was tested with the Mantel test for Dissimilarity Matrices with the Pearson correlation method. For both alpha and beta diversity, no significant difference (p values > 0.1) were found for the data set rarefied to 8709 sequences, the lowest number of sequences retrieved from any sample, so all samples were rarefied to that level (Supporting Information Fig. S1A,B), despite losing some diversity from a few samples with very abundant reads (Supporting Information Fig. S1C).

Diversity metrics

Separate from the tests of the effects of rarefaction on alpha and beta diversity of the whole data set ($n = 282$ samples for community structure analysis), the alpha diversity and beta diversity of each sample was examined in the context of oceanographic parameters and ASV relative read abundance. Toward alpha diversity, the Shannon diversity index was calculated for each sample of the rarefied data set using the *vegan* function *diversity*. The Shannon diversity index of groups of samples from different stations and depths were compared using ANOVA implemented in R with the function *anova* and Pearson's analysis was used to test the strength and significance of the correlation between alpha diversity and depth. Toward beta diversity, Bray Curtis dissimilarities were calculated and visualized via nonmetric multidimensional scaling (NMDS) using the *vegan* function *metaMDS*. To test for differences between subsets of samples, analysis of similarity (ANOSIM),

vegan function *anosim*, was used on the dissimilarity matrices. The vegan function *envfit* was used to fit environmental vectors onto the NMDS ordinations and report the goodness of fit statistic (r) and p values for each environmental variable.

ASV phylogenies

MEGA X (Stecher et al. 2020) was used for phylogenetic analysis of ASVs. Phylogenies of ASVs and reference ITS sequences were constructed by first aligning sequences with CLUSTALW (Thompson et al. 1994) and trimming all sequences to 234 base pairs. Maximum likelihood trees with the model (+G) were calculated with 1000 bootstraps. Trees were visualized and annotated in iTOL v4 (Letunic and Bork 2019).

Hierarchical clustering and Pearson's correlations

ASVs were clustered hierarchically based on their percent relative abundance across three 24-depth profiles (one from each station). First, the distance matrix between ASVs was computed using the Euclidean method with the R function *dist*. Hierarchical cluster analysis on the distances was computed using the R function *hclust* and the average linkage method. Clustering data were extracted and visualized using the R package *ggdendro*. Correlations between each ASV were determined using the function *corr.test* in the R package *psych* using the Pearson method with Holm-Bonferroni adjustments for multiple tests. Pearson correlations were reported for

pairwise comparisons with p values less than 0.01. Due to the drawbacks of correlation analysis with compositional data, we present correlations alongside hierarchical clustering and the relative abundance data for increased context. Groups of ASVs with strong correlations were determined by eye and groups defined via hierarchical clustering. Comparisons of distances between ASVs determined via hierarchical clustering vs. phylogenetic analysis were done with the Mantel test implemented using the function *PhyloMantel* from the R package *evolq* (Melo et al. 2015).

Variance partitioning analysis

To estimate the contributions of environmental variables to patterns of ASV abundance and distribution, variance partitioning analysis was performed using the function *calc.relimp* from the R package *relaimpo* (Groemping 2006). Specifically, analysis used the *lmp* metric (R^2 partitioned by averaging over orders, Lindemann, Merenda, and Gold) and the relative importances of all tested variables were summed to 100% for each ASV. Environmental variables tested were limited to those with p values less than 0.05 from vegan *envfit* analysis and included: temperature, depth, N + N, phosphate, silicate, mixed layer depth, latitude, and oxygen. Overall R^2 of variance partitioning results are available in Supporting Information Table S2.

Results

Oceanographic gradients across the North Pacific Subtropical Front

Across the North Pacific Subtropical Front, MODIS satellite observations showed large-scale gradients in sea surface temperature from 22°C at Sta. 26N to 19°C at Sta. 30N to 16°C at 35N (Fig. 1). Temperature data gathered from replicate CTD casts agreed with satellite data (Fig. 2A–C) and revealed distinct water column structure for each station. At 26N, there was no clear surface mixed layer (Fig. 2C). Salinity and temperature profiles indicated stratification from the surface to 145 m. At 30N, casts were near the main jet of the North Pacific Subtropical Front and showed a mixed layer ranging from 25 to 50 m (Fig. 2C). An increase in winds between the first and last cast at the 30N station homogenized the mixed layer but little deepening was observed, likely due to the stability of existing stratification. Depth profiles at 35N exhibited the deepest mixed layers (60–90 m deep), with abrupt transitions to the thermocline, likely caused by recent higher winds and strong existing stratification. Nutrient concentrations in the thermocline (100–150 m) increased with latitude, but only silicate increased with latitude in the mixed layer (Fig. 2A–C, Supporting Information Fig. S2). We examined the positioning of the three stations relative to temperature and salinity transitions that define the subtropical domain, subtropical front, and subtropical to subarctic transition zone according to Roden (1991) (Supporting Information Fig. S3; Fig. 2,

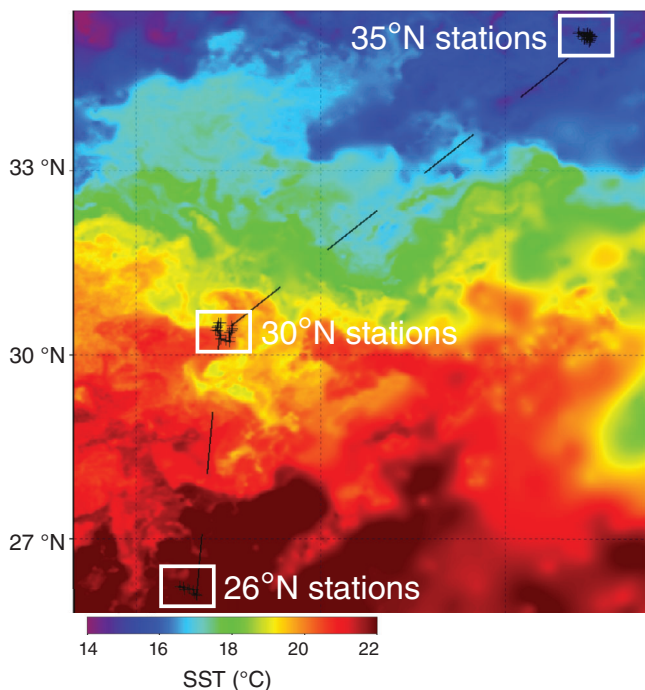


Fig 1. MODIS sea surface temperature from the cruise period (March 2017) across the North Pacific Subtropical Front. Each white box is the location of one high-resolution sampling effort. Individual depth profiles are shown as black crosses.

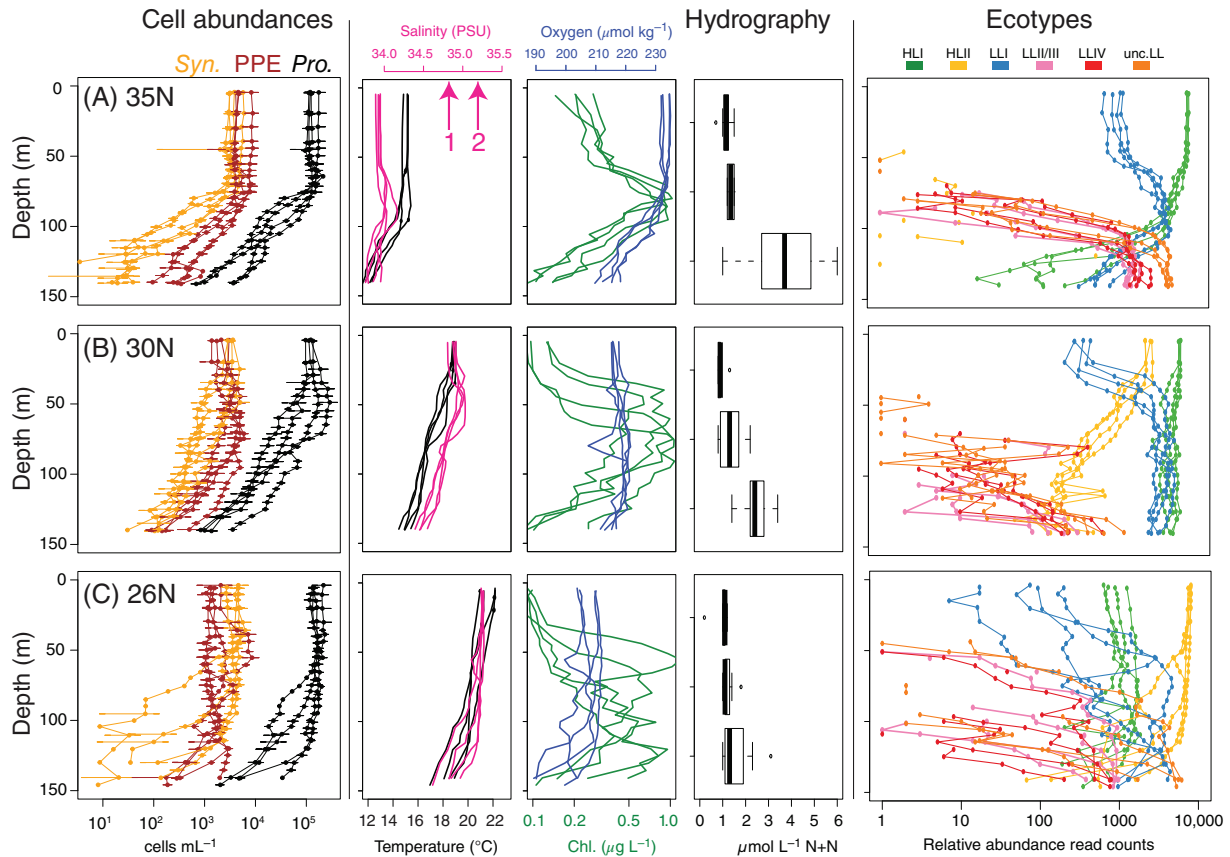


Fig 2. For each station, measurements of phytoplankton cell abundances (left), hydrographic parameters (middle three panels), and *Prochlorococcus* ecotype-level community structure (right) over depth, organized by latitude in correspondence to Fig. 1. Chl., chlorophyll, N + N, NO₂ + NO₃; PPEs, pigmented picocaryotes; Pro., *Prochlorococcus*; Syn., *Synechococcus*. Across all panels, lines are the results from individual casts at each station. The northern (1) and southern (2) boundaries of the subtropical front are indicated by pink arrows at salinities 34.8 and 35.2 according to Roden (1991).

salinity panel). 26N stations were within the subtropical domain, south of the southern boundary of the subtropical front as indicated by the 35.2 isohaline. 30N stations were within the subtropical front, between the 35.2 and 34.8 isohalines. 35N stations were north of the northern boundary of the subtropical front defined by the 34.8 isohaline.

***Prochlorococcus* dominates the phytoplankton community across gradients over latitude and depth**

Across the oceanographic gradients of the North Pacific Subtropical Front, *Prochlorococcus* numerically dominated the phytoplankton community (Fig. 2A–C and Supporting Information Fig. S4), similar to previous studies across the frontal boundary (Ribalet et al. 2015). In surface waters, *Prochlorococcus* mean abundances varied by 1.5-fold (Fig. 2, Supporting Information Fig. S4) over an 8°C shift in surface temperatures. Over depth, *Prochlorococcus* abundances ranged from a maximum of 2×10^5 cells mL⁻¹ in the surface to a minimum of 10^3 cells mL⁻¹ at 145 m across the three stations. *Synechococcus* and PPE populations were two orders of

magnitude less abundant than *Prochlorococcus* across depths (Fig. 2A–C).

Sampling at 5 m intervals over depth allowed us to examine how fine-scale physical structure of the water column corresponded to *Prochlorococcus* abundances. Depth profiles of *Prochlorococcus* abundance were distinct at each station, with transitions in abundance mirroring transitions in the density structure of each water column (i.e., stratification, mixing, and density anomalies) (Fig. 2A–C). At 26N, *Prochlorococcus* abundances were constant over depth until 75 m, then declined gradually to remain at over 100,000 cells mL⁻¹ at the base of the euphotic zone (145 m), reflecting the absence of a mixed layer and weak stratification. At 30N, *Prochlorococcus* abundances were highest at depths just below the surface mixed layer, possibly benefiting from abundant light and nutrients in the thermocline below the shallow mixed layer. At 35N, *Prochlorococcus* abundances were constant over depth through the mixed layer, increased slightly below the mixed-layer boundary, then declined sharply with depth, reflecting the deep mixed layer and sharp boundary to the thermocline.

Ecotype-level community structure across the North Pacific Subtropical Front

Transitions in *Prochlorococcus* community structure at the ecotype-level mirrored transitions in physical structure of the water column and *Prochlorococcus* abundance patterns over depth (Fig. 2A–C, Supporting Information Fig. S5), as measured by relative read abundances of ASVs belonging to each ecotype. Relative read abundance of HLI and HLII clades were highest in the surface and shifted from HLII-dominated at Sta. 26N to HLI-dominated at Sta. 30N and 35N. Other HL ecotypes were not detected.

The highest abundances of LL ecotypes occurred at Sta. 35N, where the strongest stratification coincided with abundant LL populations in the thermocline. Relative read abundance of LLI was most abundant below the mixed layer at each station. Other low light ecotypes (LLII-III, LLIV, and unclassified low light sequences) were most abundant at the deepest depths, and in some cases were not detected. Relative read abundance for the LLI clade was higher at Sta. 30N and 35N, consistent with a latitudinal transect in the Atlantic (Johnson et al. 2006).

Major shifts in microdiversity underlie *Prochlorococcus* ecotype and abundance patterns

In order to understand the microdiversity underlying ecotype-level patterns in *Prochlorococcus* community structure, and patterns in abundance over depth and latitude, we examined patterns of alpha and beta diversity within and between samples based on 4100 ASVs identified as *Prochlorococcus*. In surface waters (above the mixed layer depth for all stations, less than 50 m), alpha diversity was dramatically different across stations (ANOVA, p values < 0.01 , Fig. 3A, Supporting Information Fig. S6A), with highest diversity at Sta. 26N and lowest diversity at the farthest North Sta. 35N. At intermediate depths 50–100 m, 26N alpha diversity was higher than both 30N and 35N (ANOVA, p values < 0.01); however, 30N and 35N were not significantly different. At the deepest depths (greater than 100 m), 26N samples were more diverse than 30N samples (ANOVA, $p < 0.01$) but not statistically different from 35N samples ($p = 0.68$). The diversity in samples from 35N spanned the greatest range over depth, from the lowest of the three stations at the surface to among the highest at the deepest depths (Fig. 3A). For 35N and 26N, alpha diversity increased significantly over depth (Pearson's correlation analysis, $p < 0.01$), while alpha diversity did not increase significantly over depth at 30N ($p = 0.19$) (Supporting Information Fig. S6B).

To test similarity of samples with respect to *Prochlorococcus* ASV community structure, we calculated beta diversity for all sample pairs using Bray Curtis dissimilarity and used NMDS to visualize relationships (Fig. 3B). The NMDS analysis showed that the three stations were significantly distinct in *Prochlorococcus* ASV-level community structure (Fig. 3B, ANOSIM significance = 0.0001). Within each group of samples

from the same station, samples also grouped together by depth.

Environmental variables were examined for the strength (R^2) and significance (p value < 0.01) of their correlation to the community structure at each station. Correlating environmental variables to the NMDS ordination axes revealed that temperature, oxygen, and latitude most strongly influenced ASV-level community structure ($R^2 > 0.8$). Other significantly correlated variables included depth, mixed layer depth, silicate, N + N, and phosphate, but for these variables R^2 ranged from 0.324 to 0.645. Chlorophyll, photosynthetically active radiation, ammonium (NH_4^+), and nitrite (NO_2^-) were not significantly correlated with community structure.

High diversity in low-light adapted clades

In order to identify ecologically relevant patterns of subecotype-level diversity, we examined patterns of ASV abundance, distribution, and correlation to environmental parameters. Of the 4230 ASVs identified with more than 100 sequence reads, 4100 belonged to *Prochlorococcus* and 130 belonged to *Synechococcus*. Of the *Prochlorococcus* ASVs, most belonged to LL clades (Fig. 4). Most *Prochlorococcus* ASVs best matched reference sequences from the LLI clade ($n = 1171$) or phylogenetically unclassified low light reference sequences from environmental samples ($n = 1924$), consistent with the high diversity known among LL ecotypes (Rocap et al. 2002; Ahlgren and Rocap 2006; Lavin et al. 2008). There were 23 ASVs belonging to the LLIV clade and 249 belonging to LLII-III clades. Sequences matching LLV and LLVI ecotypes were detected however at percent similarity lower than 90%, thus were included in the pool of ASVs categorized as unclassified low light. ASVs belonging to HL clades best matched HLI and HLII ecotypes and each contained fewer than 500 ASVs. No ASVs matching HLIV, HLVI, or HNLC ecotypes were detected.

Two ASVs are ubiquitous across the North Pacific Subtropical Front

To further examine community structure among *Prochlorococcus* ASVs, the frequency of each ASV across samples was compared to its maximum relative abundance and median abundance where present (Fig. 4) in an approach similar to one used to test connectivity across the Great Lakes (Paver et al. 2020). Several groups of ASVs were noted, each with distinct relationships between frequency and abundance. The first group of ASVs ($n = 2$) were ubiquitous and dominant, including a HLI and a LLI ASV. These two ASVs, ASV1 and ASV2, were present in nearly 100% of samples, and at greater than 25% of the community when present, across the strong oceanographic gradients of the study region. Furthermore, ASV1 comprised a large proportion (74%) of all HLI sequence reads. ASV2 comprised 38% of all LLI sequence reads.

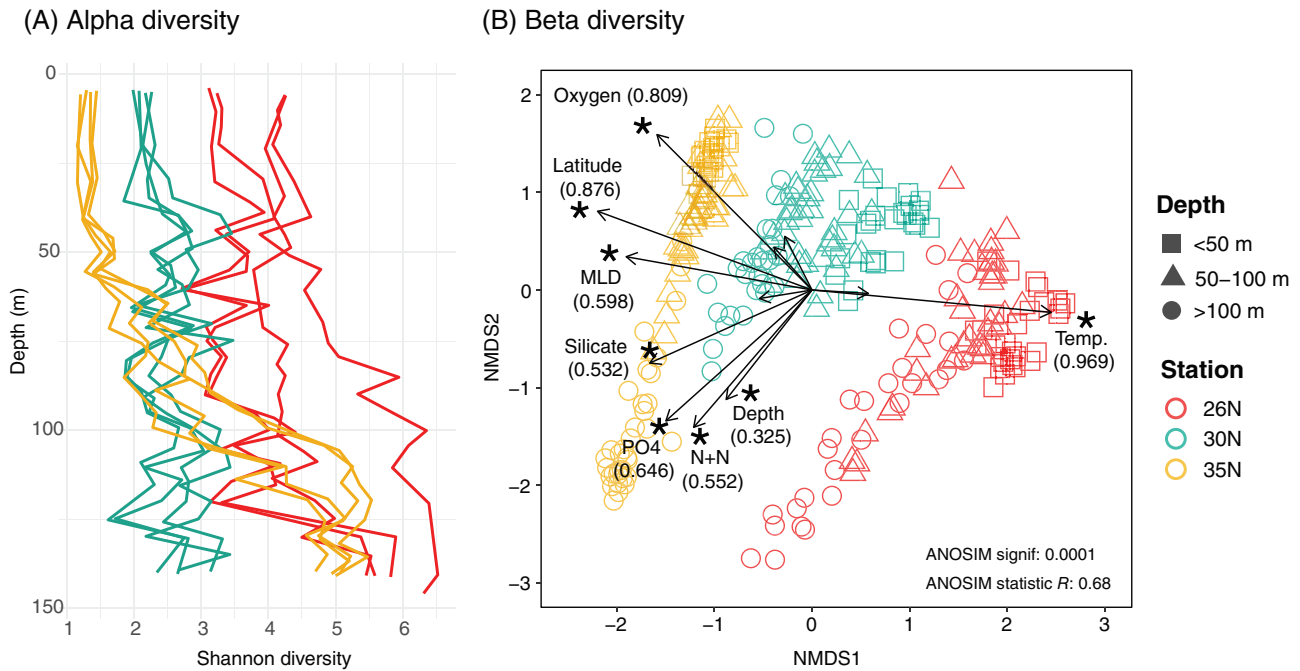


Fig 3. Diversity of *Prochlorococcus* within and between samples based on 4100 *Prochlorococcus* ASVs. **(A)** Shannon diversity over depth shows higher diversity in surface waters from 26N and lowest diversity in surface waters from 35N. Over depth, diversity increased significantly at 26N and 35N but was constant at 30N (Supporting Information Fig. S6). **(B)** NMDS calculated on Bray-Curtis dissimilarities between all samples. Symbols are colored by station and shape indicates depth. Significance of clustering by station is reported for ANOSIM. Environmental drivers that were tested for correlation to the NMDS patterns are indicated as black vectors. Drivers with p values < 0.01 are indicated by asterisks (*) and the R^2 is presented in parentheses for significant variables. Nonsignificant variables are not labeled and included: PAR, NO_2^- , NH_4^+ , and fluorescence.

Based on the distribution of ASVs on the plot of the ubiquity (frequency of presence across samples) vs. abundance (maximum abundance) (Fig. 4), we categorized ASVs into additional groups. The second group comprised ASVs that were frequent (found in 50–90% of samples) and abundant (maximum abundance of 0.5–25%) and were composed of HLI, HLII, and LLI ASVs ($n = 13$). A third group of ASVs were abundant (0.5–25% maximum abundance) but present only occasionally (5–50% of the samples), and included representatives from all detected ecotypes ($n = 206$). These three groups comprise the 221 most abundant and frequent ASVs, which were examined further. Most ASVs were part of a fourth major group that was rarely present in samples (< 5% of samples) and always present as minor community members (< 0.5% of the community). Most ASVs (94.6%) fell into this category ($n = 3879$).

Subecotypes driven by distinct environmental factors

The ecological relevance of *Prochlorococcus* microdiversity was examined by grouping the most abundant and frequent ASVs (defined above, Fig. 4) based on phylogeny, or hierarchical clustering on relative abundance patterns, testing ASV correlations across depths and stations, then identifying the relative contributions of environmental variables to observed distribution patterns with variance partitioning analysis.

For LLI ASVs, hierarchical clustering based on relative abundance patterns and phylogenetic analysis based on ASV sequences revealed distinct groups of ASVs (Fig. 5A, Supporting Information Fig. S7A). Several clusters of ASVs defined by hierarchical clustering displayed distinct shared patterns across stations and depths as revealed in plots of percent maximum abundance at three representative casts across all three stations and high R^2 value for significant ($p < 0.01$, adjusted for multiple comparisons) Pearson's correlations among ASVs (Fig. 5A, heat map). LLI ASVs formed at least four major clusters with strong and significant correlations, labeled as LLI-a through LLI-d. LLI-a, LLI-c, and LLI-d were most abundant at depths below the mixed layer but above 120 m. These clusters exhibited different patterns of latitudinal partitioning. LLI-a was most abundant in higher latitudes (30N and 35N), while LLI-c was most abundant at 26N. LLI-d was broadly distributed across all stations, and included the dominant and ubiquitous LLI ASV2 (Fig. 4). In contrast to the other LLI clusters, LLI-b was present primarily at 35N and deeper in the water column than the other LLI ASV clusters. The depth distribution of LLI-b was more similar to the other LL ecotypes (Fig. 5B) than to LLI. To confirm the identity of LLI-b ASVs as LLI rather than another LL clade, a separate BLAST analysis confirmed best hits to LLI strains (e.g., *Prochlorococcus* PAC1, NATL2A, and MIT0917) at an average of 97.1% similarity, with stronger average sequence similarity (98.4%) to

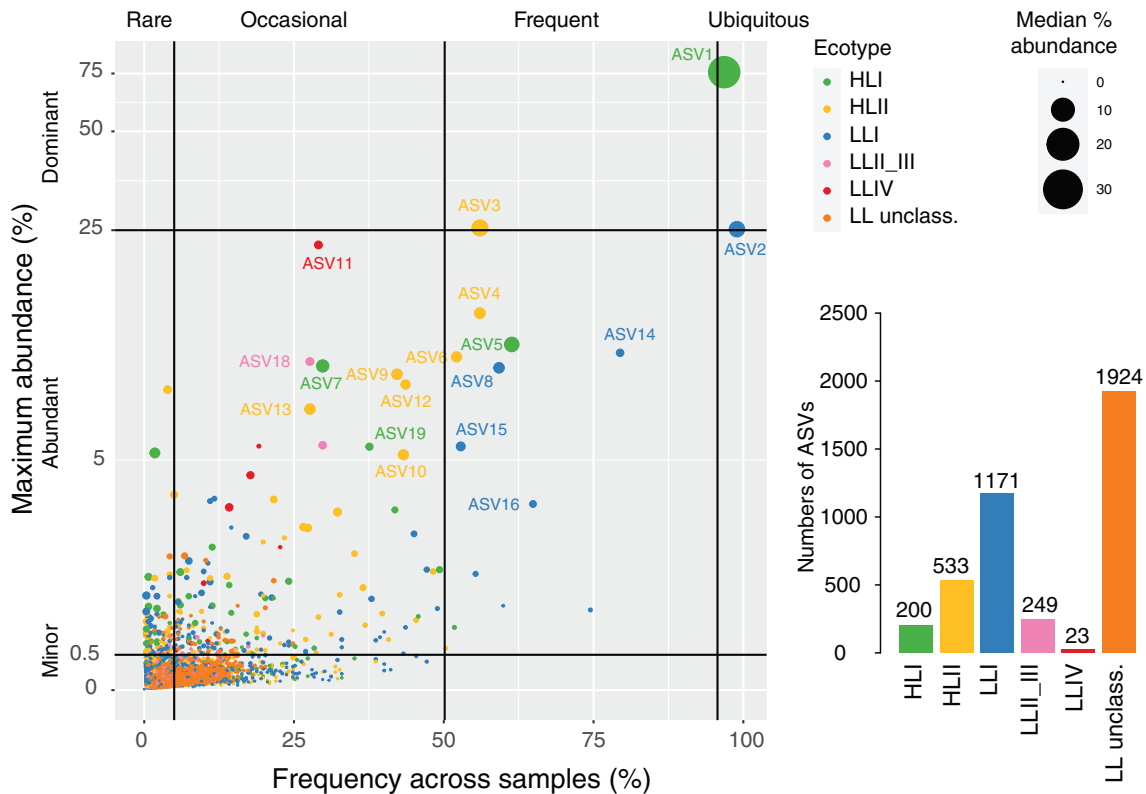


Fig 4. The numbers of ASVs belonging to each *Prochlorococcus* ecotype (bar plot) and comparison of the frequency of ASVs across samples as a percent of total samples (*x*-axis), maximum abundance relative to the rest of the community in one sample (*y*-axis), and median abundance where present (size symbol, legend) (scatter plot). Colors show ecotype identity. ASVs were classified into several groups determined by their position on the plot defined by ranges of frequency and maximum abundance (i.e., “Minor,” “Abundant,” “Dominant,” “Rare,” “Occasional,” “Frequent,” “Ubiquitous”).

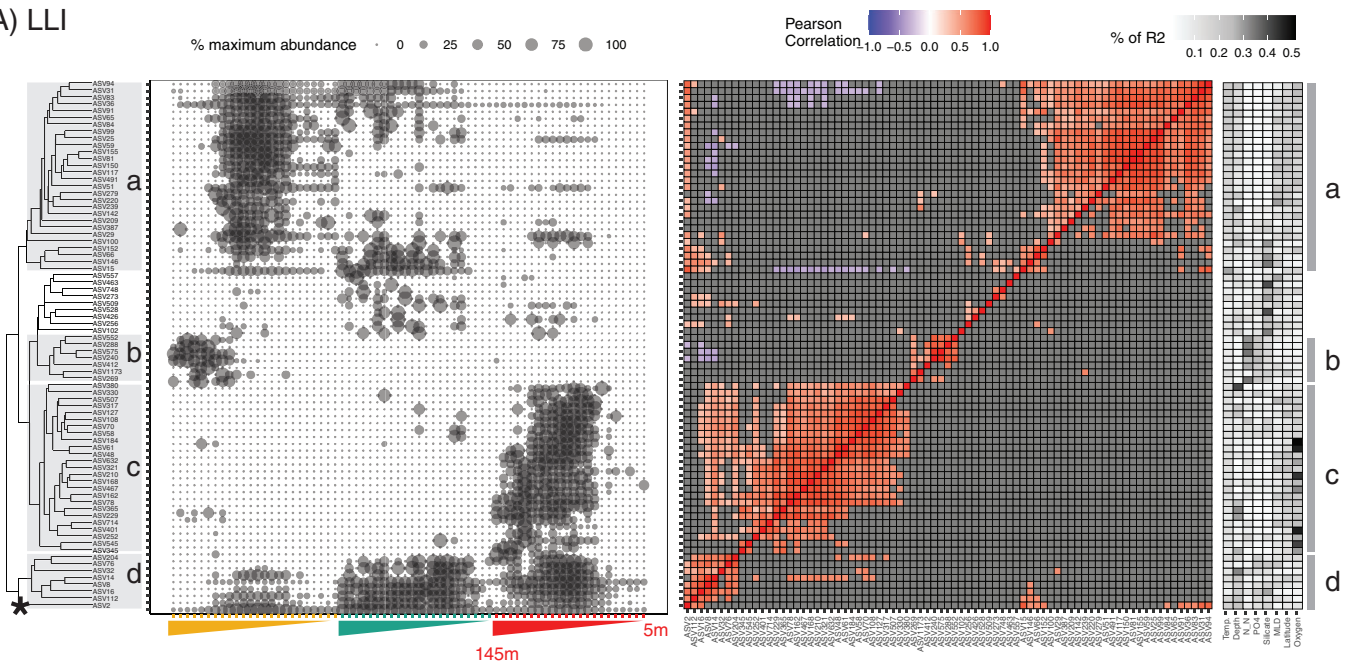
uncultured environmental *Prochlorococcus* of unknown phylogeny (Supporting Information Table S1). From variance partitioning analysis, environmental variables were relatively even in their contribution to explaining abundance patterns of LLI-a and LLI-d. In contrast, oxygen contributed to explaining LLI-c patterns and nutrient concentrations (N + N and PO₄) contributed most to explaining LLI-b patterns. Together the variance partitioning results reflect the corresponding patterns of these parameters and ASV clusters with latitude (Fig. 2). For all ASVs, the summed contribution of all environmental variables was 40%, suggesting that additional nonmeasured variables may contribute to the patterns observed. These could include organic nutrients (Muñoz-Marín et al. 2020) or top-down pressure from grazers (Frias-lopez et al. 2009; Apple et al. 2011) and cyanophage (Ahlgren et al. 2019). Several ASVs did not fall into these four large clusters and were only significantly correlated in abundance to a few or to no other ASVs. Most ASVs in this category were sparsely present across stations and at low abundance and frequency (rare and/or minor in Fig. 4).

For LL ecotypes other than LLI (Other LL), three major clusters of ASVs emerged from hierarchical clustering, assessment of abundance across representative depth profiles, strength of

correlation of the ASV patterns, and variance partitioning analysis (Fig. 5B). LLother-a was abundant at Sta. 35N and distribution patterns were most explained by nutrient concentrations in variance partitioning analysis. In direct contrast, LLother-c was abundant at Sta. 26N and variance patterns were best explained by oxygen. LLother-b was present at both Sta. 26N and 35N, with variance partitioning analysis results similar to LLother-a with nutrient concentrations best explaining variation. As noted above, these variance partitioning analysis results reflect how oxygen and nutrient levels vary between 26N and 35N (Fig. 2). Many ASVs were correspondingly only present at only station or the other, but there were several ASVs in both clusters that were present at appreciable levels at both stations. These three clusters were composed of ASVs from multiple different LL ecotypes (LLII-III, LLIV, and unclassified LL sequences) and there was no coherent latitudinal partitioning of taxa at the broader ecotype level (Fig. 2, Supporting Information Fig. S5). Unlike LLI, all of the non-LLI ASVs (Fig. 1, Supporting Information Fig. S4) shared a common pattern in that they were generally absent from surface waters and peaked at the deepest depths sampled (145 m).

The two HL ecotypes differed greatly in how their ASVs were divided into distinct clusters based on abundance

(A) LLI



(B) Other LL

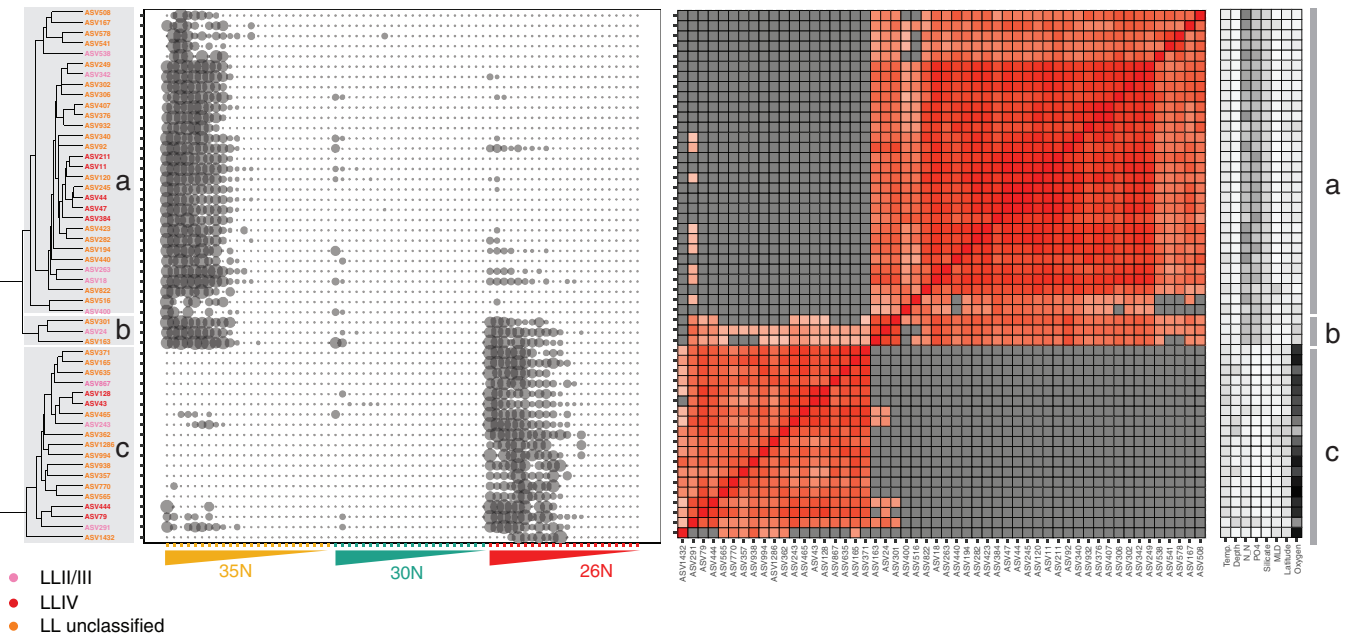


Fig 5. Low light (LL) ASV clusters by hierarchical clustering (dendrogram, left), distribution patterns over depth and station (left middle), strength of significant (p value < 0.01) correlations of ASVs abundance patterns (heat map), and variance partitioning analysis to determine the relative importance of environmental parameters (% R^2 , right) for ASVs best matching LLI (A) and non-LLI LL clades (“other” LL) (B). LL ecotypes in (B) are identified at the ecotype level by coloring of the ASV name. Symbol size represents the percent maximum abundance for each ASV for three representative stations (yellow, green, red) and over depth (indicated by the colored wedges). Specific clusters of ASVs that are discussed in the text are defined by gray boxes and labeled with black text on either side of the plots. The asterisk (*) on (A) marks one of the two ubiquitous ASVs defined in Fig. 4. Overall R^2 of variance partitioning results are available in Supporting Information Table S2.

patterns and variance partitioning analysis (Fig. 6). Examining hierarchical clustering and correlations determined by Pearson’s correlation analysis, HLI ASVs formed four major hierarchical clusters labeled HLI-a through HLI-d that were

partitioned by station (Fig. 6A), similar to LLI ASV clusters (Fig. 5A). Cluster HLI-a was abundant in surface waters of 35N, while HLI-b was abundant in 35N and 30N surface waters. The most ubiquitous and abundant ASV, ASV1 (Fig. 4),

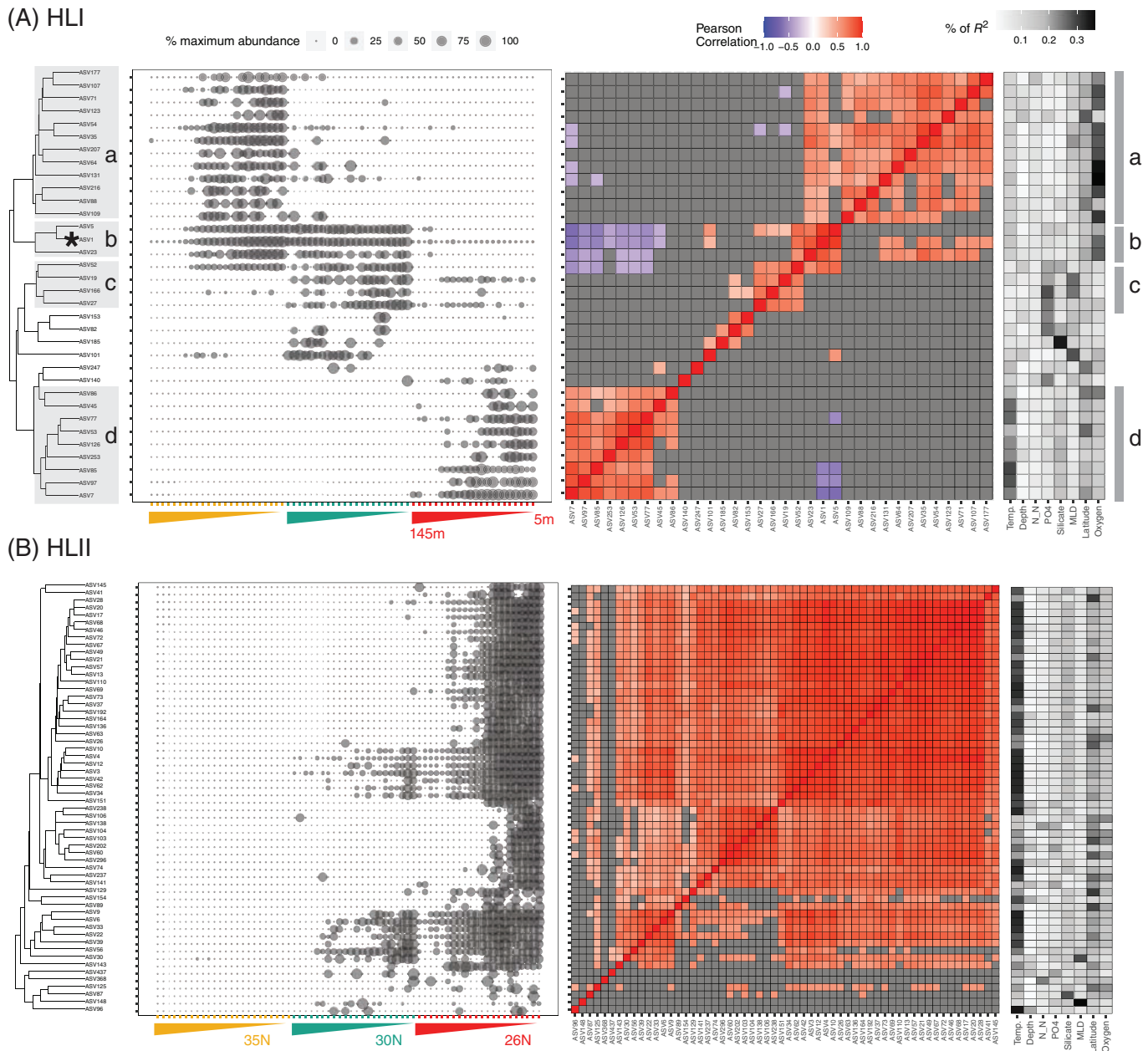


Fig 6. High light (HL) ASV clusters by hierarchical clustering (dendrogram, left), distribution patterns over depth and station (left middle), strength of significant (p value < 0.01) correlations of ASVs abundance and distribution patterns (heat map), and variance partitioning analysis to determine the relative importance of environmental parameters ($\%R^2$, right) for ASVs best matching HLI (A) and HLII (B) clades. Symbol size represents the percent maximum abundance for each ASV for three representative stations (yellow, green, red) and over depth (indicated by the colored wedges). Specific clusters of ASVs that are discussed in the text are defined by gray boxes and labeled with black text on either side of the plots. The asterisk (*) on (A) marks one of the two ubiquitous ASVs defined in Fig. 4. Overall R^2 of variance partitioning results are available in Supporting Information Table S2.

belonged HLI-b. HLI-c was most abundant at Sta. 30N but some of its ASVs were found appreciably at the other stations. HLI-d was abundant almost exclusively at Sta. 26N. Of the environmental variables tested, correlations to oxygen explained most of the HLI-a and HLI-b patterns while temperature correlations explained most of the HLI-d patterns. Phosphate best explained HLI-c patterns. These variance partitioning

results were consistent with the corresponding latitudinal patterns of these parameters and ASV clusters. HLI ASVs that were not assigned a cluster were present at low or sporadic abundance and frequency.

In contrast to HLI, most HLII ASVs formed one major cluster according to Pearson's correlation analysis of abundance and distribution over depth (Fig. 6B). For HLII, temperature

contributed the most of the tested environmental variables to the observed distribution patterns (Fig. 6B). Most HLII ASVs were only abundant at 26N, but we did observe two smaller clusters of ASVs that were present at 30N as well. As with LLI, there were a handful of ASVs that were only significantly correlated to a few or no other ASVs and belonged to the rare, low abundance category in Fig. 4.

For each of the ecotypes analyzed above (LLI, other LL ecotypes, HLI, and HLII), we assessed the congruence of ASV relatedness by abundance-based hierarchical clustering vs. phylogeny. We did not find cohesive distribution patterns among phylogenetically defined ASV groups (Supporting Information Figs. S7, S8). While the short region of the ITS (~ 234 bp) allows accurate assignment of sequences at the coarse ecotype level, confirmed by maximum likelihood analysis relative to reference sequences with well-characterized phylogeny (Supporting Information Fig. S9), the ITS can be difficult to align due to frequent insertions and deletions (Rocap et al. 2002). As a result, phylogenetic analysis of ASVs within ecotypes produced low bootstrap support at most nodes (Supporting Information Fig. S9). Likewise, high diversity among LL *Prochlorococcus*, and their deep branching phylogenies (Huang et al. 2011), mean that appropriate reference sequences may not exist for some of the LL ASVs recovered. Combined with limitations in using BLAST percent identity to make subecotype-level assignments, it was difficult to make strong inferences of their subecotype-level genetic relatedness. In addition, we found low strength of correlation ($R^2 < 0.16$, Mantel test) between distance matrices based on phylogeny or hierarchical clustering (Supporting Information Fig. S10). Thus, we cannot rule out phylogeny as a factor in organizing ASVs that have similar abundance and distribution patterns with respect to the oceanic gradients tested here.

Discussion

Diversity and niche partitioning within low light *Prochlorococcus* ecotypes

This study is the first to reveal extensive subecotype diversity and niche partitioning of subecotype-level groups within low-light adapted *Prochlorococcus* ecotypes across oceanographic gradients in the North Pacific. Of over 4000 ASVs detected in this study, 81% belonged to LL ecotypes, with highest diversity in the LLI clade and unclassified LL clades. This is consistent with previous clone library studies of representative strain marker genes and comparative genomic analysis that show higher phylogenetic diversity among LL ecotypes than their HL counterparts (Kettler et al. 2007; Martiny et al. 2009b; Huang et al. 2011). Studies of sequenced genomes of LL *Prochlorococcus* cultured isolates (Kettler et al. 2007; Biller et al. 2014; Berube et al. 2015), analysis of full length 16S-23S rRNA ITS from environmental sequences (Rocap et al. 2002; Huang et al. 2011), and growth physiology in response to nutrients (Berube et al. 2015) and light

(Malmstrom et al. 2010) have also suggested high ecological variability among LL clades. Thus, this work links what has been learned of LL strains from studies of cultured representatives to wild ocean populations and reveals the importance of latitudinal gradients in LL niche partitioning at the subecotype level. In addition, this work also expands to the LL *Prochlorococcus* ecotypes the understanding that ecologically distinct taxa occur within coherent ecotypes of picocyanobacteria as shown previously for *Synechococcus* (Pittera et al. 2014; Farrant et al. 2016; Ahlgren et al. 2019) and HL *Prochlorococcus* (Larkin et al. 2016).

This study provides insight into the characteristics and potential environmental drivers of the LLI ecotype and the subecotype groups within it. LLI ASVs formed four major significant and cohesive clusters with respect to abundance and distribution across latitude (Fig. 5A). All but one of these four LLI clusters were most abundant at intermediate depths within the thermocline. This depth preference is consistent with previous measures of LLI abundance using qPCR primers informed from representative sequences (Ahlgren et al. 2006; Johnson et al. 2006; Malmstrom et al. 2010). While most LLI ASVs share this commonality in their distribution relative to the physical structure of the water column, the individual LLI clusters exhibited distinct geographic patterns and differences in correlations to environmental parameters (Fig. 5A). Thus, in addition to a shared trait with respect to light physiology, the broad biogeographic distribution of this ecotype may be shaped by more complex, underlying niche partitioning at the subecotype level. Such niche partitioning among subecotypes has been shown previously for HL subecotypes (Larkin et al. 2016, 2019), but was unknown for LL subecotypes. While total abundances of LLI were previously shown to vary with latitude (Johnson et al. 2006), this study reveals additional latitudinal partitioning of LLI at the subecotype level. The lower euphotic zone that LLI inhabits is largely buffered from strong temperature differences that drive marked latitudinal partitioning of HLI and HLII surface ecotypes (Johnson et al. 2006; Larkin et al. 2016). Our study, however, highlights how deep-water processes correlated to latitudinal changes in nutrient availability and oxygen levels may control the distribution of LLI populations.

A single LLI cluster peaked deep within the euphotic zone (LLI-b) and was more similar in depth distribution to other LL subecotypes such as LLIV and LLII/III (Fig. 5B) than to LLI (Fig. 5A). While these ASVs share 97–97% sequence similarity to representative LLI strains, they may represent a novel lineage of LL *Prochlorococcus* phylogenetically similar, but physiologically distinct, from LLI.

Niche partitioning of ASVs within other LL ecotypes was similar to patterns observed for LLI with very strong partitioning across stations associated with latitudinal changes in oxygen and macronutrients (Fig. 5B). These clear shifts in LL ASVs across stations are in stark contrast to the absence of strong latitudinal partitioning at the broader ecotype level (for abundant non-LLI ecotypes, LLII-III vs. LLIV) (Fig. 2).

Analysis of ASVs within HL ecotypes (Fig. 6) was generally consistent with the previous works on niche partitioning among HL subecotypes (Larkin et al. 2016). HLI ASVs formed at least four distinct clusters, that were always most abundant in surface waters, but varied in their distribution across the stations, with varied contributions of temperature, phosphate, and oxygen (Fig. 6B). In contrast, HLII ASVs formed one large cluster, strongly correlated to temperature (Fig. 6B), which contrasts with previous work that has shown strong evidence for niche partitioning within the HLII ecotype (Larkin et al. 2016, 2019). The difference between our work and other studies with regard to HLII may be explained by the presence of HLII ASVs at only one of our three stations while the previous studies analyzed a wider geographic range of samples in subtropical and tropical locations.

Our ability to address whether subecotype-level clusters of LL clades were rooted in phylogenetic similarity was limited due to limitations in the phylogenetic analysis of the short ITS reads. Overall, analyses where ASVs were clustered by sequence similarity (Supporting Information Figs. S7, S8) did not show cohesive ecological properties as those organized by hierarchical clustering (Figs. 5, 6). However, one exception was for LLIV ASVs, which demonstrated strong bootstrap support for assignment to the LLIV clade (Supporting Information Fig. S9). These robustly LLIV ASVs fell into two hierarchical clusters one present at Sta. 35N (LL-other a) and the other present at Sta. 26N (LL-other c) (Fig. 5). In previous work, HL clade subecotype-level niche partitioning has been shown to be rooted in phylogenetic similarity (Larkin et al. 2016). Thus, we attribute the lack of evidence for phylogenetic similarity among the clusters in part due to the difficulty of conducting robust phylogenetic analysis with short amplicons (i.e., low bootstrap support at nodes) that are subject to frequent substitutions and deletions. Similar time-series analysis of *Prochlorococcus* and *Synechococcus* populations did, however, find some coherence between ASVs clustered by temporal abundance patterns and phylogenetic relatedness (Ahlgren et al. 2019).

Structure and connectivity within the *Prochlorococcus* collective across a front

Of the 4100 *Prochlorococcus* ASVs detected in this study, we discovered two individual ASVs that were ubiquitous (i.e., present in nearly all samples) and dominant (more than 25% of the community) (Fig. 4). Ahlgren et al. (2019) similarly observed that a single picocyanobacterial ASV could consistently comprise 10–40% of an ecotype over time. The two ubiquitous ASVs belonged to the HLI and LLI clades, which vary significantly in functional gene content but share a core genome (Kettler et al. 2007; Biller et al. 2014, 2015). On the other hand, the majority of *Prochlorococcus* ASVs were rare and low abundance, and there were several ASVs in each ecotype that did not belong to larger identified groups with cohesive abundance patterns. This demonstrates that the majority of

Prochlorococcus diversity is composed of the long tail of low-abundance taxa (i.e., the “rare biosphere”), a commonly observed theme for marine microbes (Galand et al. 2009; Lynch and Neufeld 2015; Jousset et al. 2017). These observations also raised questions about the mechanisms behind assembly of *Prochlorococcus* communities, especially within the highly diverse LLI ecotype. At the broad level, these results highlight that the *Prochlorococcus* is composed of many distinct taxa with several different life strategies.

Previous high resolution transects have documented sharp biological transitions across the North Pacific Subtropical Front including in the abundance and diversity of nitrogen fixing cyanobacteria (Gradoville et al. 2020) and *Prochlorococcus* growth rates (Ribalet et al. 2015). From these data, it has been unclear whether gradients in this region shape communities through environmental selection or whether the North Pacific Subtropical Front geographically partitions and isolates cyanobacterial populations. The presence of multiple LLI ASV clusters is consistent with their ability to thrive in such transitional regions with fluctuating conditions, perhaps analogous to *Synechococcus* ecotypes (i.e., clades XV and XVI) that appear to preferentially occur in similar ecotone sites at the junctions of major water masses (Sohm et al. 2015). Conversely, non-LLI ASVs exhibit sharp partitioning across the three stations, with many ASVs absent or at low relative abundance in the North Pacific Subtropical Front (30N), suggesting they are better adapted to more stable water masses.

Despite strong partitioning across the front in community structure, the presence of two ubiquitous and dominant *Prochlorococcus* ASVs throughout, and on either side, of this oceanographic frontal region suggests connectivity across the front. Such connectivity is supported by Lagrangian particle tracking models that predict low minimum connection times in the North Pacific (less than 5 yr) (Jönsson and Watson 2016). Thus, we attribute the distinct *Prochlorococcus* diversity patterns (Fig. 3A) and community structure (Fig. 3B) documented across this region to the importance of environmental selection in the diversification of *Prochlorococcus* rather than geographical isolation. This observation is consistent with the important role of environmental selection, rather than dispersal limitations, in driving community structure of another globally dominant phytoplankton group, diatoms (Whittaker and Rynearson 2017), but inconsistent with a comparison of *Prochlorococcus* community structure on a much larger geographic scale, between the Pacific and Atlantic Oceans (Kashtan et al. 2017).

Conclusions

This work examines *Prochlorococcus* subecotype-level diversity over large-scale gradients in latitude and fine-scale gradients in depth across the North Pacific Subtropical Front. In particular, this work reveals that most *Prochlorococcus* diversity

resides in LL ecotypes and that subecotype level clusters of LL ASVs demonstrate distinct and strong patterns of niche partitioning, similar to what has been seen in previous work for HL ecotypes. Specifically, this work reveals more about of the ecology of the enigmatic LLI clade, which has been shown to have unique light physiology properties in culture and distinct seasonal dynamics, but is less well studied than HL counterparts in the analysis of the *Prochlorococcus* collective community structure dynamics. We also found a combination of ubiquitous and habitat specific ASVs across the North Pacific Subtropical Front, suggesting that ecosystems across the front are connected but that microbial communities on either side are shaped by environmental selection. Understanding how the shifts we observed in ecotype- and subecotype-level community structure of LL clades affects *Prochlorococcus* contributions to primary productivity, food webs, and nutrient cycling will improve prediction of the dynamics of global processes.

References

- Ahlgren, N. A., and G. Rocap. 2006. Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. *Appl. Environ. Microbiol.* **72**: 7193–7204. doi:10.1128/AEM.00358-06
- Ahlgren, N. A., G. Rocap, and S. W. Chisholm. 2006. Measurement of *Prochlorococcus* ecotypes using real-time polymerase chain reaction reveals different abundances of genotypes with similar light physiologies. *Environ. Microbiol.* **8**: 441–454. doi:10.1111/j.1462-2920.2005.00910.x
- Ahlgren, N. A., J. N. Perelman, Y.-C. Yeh, and J. A. Fuhrman. 2019. Multi-year dynamics of fine-scale marine cyanobacterial populations are more strongly explained by phage interactions than abiotic, bottom-up factors. *Environ. Microbiol.* **21**: 2948–2963. doi:10.1111/1462-2920.14687
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410. doi:10.1016/S0022-2836(05)80360-2
- Apple, J. K., S. L. Strom, B. Palenik, and B. Brahmsha. 2011. Variability in protist grazing and growth on different marine *Synechococcus* isolates. *Appl. Environ. Microbiol.* **77**: 3074–3084. doi:10.1128/AEM.02241-10
- Berube P. M., and others. 2015. Physiology and evolution of nitrate acquisition in *Prochlorococcus*. *The ISME Journal* **9**: 1195–1207. <http://dx.doi.org/10.1038/ismej.2014.211>
- Berube, P. M., and others. 2018. Single cell genomes of *Prochlorococcus*, *Synechococcus*, and sympatric microbes from diverse marine environments. *Sci. Data* **5**: 180154. doi:10.1038/sdata.2018.154
- Bibby, T. S., Mary, I., Nield, J., Partensky, F., Barber, J. 2003. Low-light-adapted *Prochlorococcus* species possess specific antennae for each photosystem. *Nature* **424**: 1051–1054. <http://dx.doi.org/10.1038/nature01933>
- Biller, S. J., and others. 2014. Genomes of diverse isolates of the marine cyanobacterium *Prochlorococcus*. *Scientific Data* **1**: (1) <http://dx.doi.org/10.1038/sdata.2014.34>
- Billar, S. J., Berube, P. M., Lindell, D., Chisholm, S. W. 2015. *Prochlorococcus*: the structure and function of collective diversity. *Nat. Rev. Microbiol.* **13**: 13–27. <http://dx.doi.org/10.1038/nrmicro3378>
- Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**: 581–583. doi:10.1038/nmeth.3869
- Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, and E. K. Costello. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**: 335–336. doi:10.1038/nmeth.f.303
- Chandler, J. W., Y. Lin, P. J. Gainer, A. F. Post, Z. I. Johnson, and E. R. Zinser. 2016. Variable but persistent coexistence of *Prochlorococcus* ecotypes along temperature gradients in the ocean's surface mixed layer. *Environ. Microbiol. Rep.* **8**: 272–284. doi:10.1111/1758-2229.12378
- Chisholm, S. W., Frankel, S. L., Goericke, R., Olson, R. J., Palenik, B., Waterbury, J. B., West-Johnsrud, L., Zettler, E. R. 1992. *Prochlorococcus marinus* nov. gen. nov. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll a and b. *Archives of Microbiology* **157**: 297–300. <http://dx.doi.org/10.1007/bf00245165>
- Farrant, G. K., and others. 2016. Delineating ecologically significant taxonomic units from global patterns of marine picocyanobacteria. *Proc. Natl. Acad. Sci. USA* **113**: E3365–E3374. doi:10.1073/pnas.1524865113
- Flombaum, P., and others. 2013. Present and future global distributions of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. USA* **110**: 9824–9829. doi:10.1073/pnas.1307701110
- Frias-lopez, J., A. Thompson, J. Waldbauer, and S. W. Chisholm. 2009. Use of stable isotope-labelled cells to identify active grazers of picocyanobacteria in ocean surface waters. *Environ. Microbiol.* **11**: 512–525. doi:10.1111/j.1462-2920.2008.01793.x
- Galand, P. E., E. O. Casamayor, D. L. Kirchman, and C. Lovejoy. 2009. Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc. Natl. Acad. Sci. USA* **106**: 22427–22432. doi:10.1073/pnas.0908284106
- Gradoville, M. R., and others. 2020. Latitudinal constraints on the abundance and activity of the cyanobacterium UCYN-A and other marine diazotrophs in the North Pacific. *Limnol. Oceanogr.* **65**: 1858–1875. doi:10.1002/lno.11423
- Groemping, U. 2006. Relative importance for linear regression in R: The package relaimpo. *J. Stat. Softw.* **17**: 1–27. doi:10.18637/jss.v017.i01
- Huang, Sijun, Wilhelm, Steven W, Harvey, H Rodger, Taylor, Karen, Jiao, Nianzhi, Chen, Feng 2012. Novel lineages of

- Prochlorococcus* and *Synechococcus* in the global oceans. *The ISME Journal* **6**: (2) 285–297. <http://dx.doi.org/10.1038/ismej.2011.106>
- Johnson, Z. I., E. R. Zinser, A. Coe, N. P. McNulty, E. Malcolm S Woodward, and S. W. Chisholm. 2006. Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311**: 1737–1740. doi:10.1126/science.1118052
- Jönsson, B. F., and J. R. Watson. 2016. The timescales of global surface-ocean connectivity. *Nat. Commun.* **7**: 11239. doi:10.1038/ncomms11239
- Jousset, A., and others. 2017. Where less may be more: How the rare biosphere pulls ecosystems strings. *ISME J.* **11**: 853–862. doi:10.1038/ismej.2016.174
- Kashtan, N., and others. 2014. Single-cell genomics reveals hundreds of coexisting subpopulations in wild *Prochlorococcus*. *Science* **344**: 416–420. doi:10.1126/science.1248575
- Kashtan, Nadav, Roggensack, Sara E, Berta-Thompson, Jessie W, Grinberg, Maor, Stepanauskas, Ramunas, Chisholm, Sallie W 2017. Fundamental differences in diversity and genomic population structure between Atlantic and Pacific *Prochlorococcus*. *The ISME Journal* **11**: 1997–2011. <http://dx.doi.org/10.1038/ismej.2017.64>
- Kelley, D. E. 2018. Oceanographic analysis with R. doi:10.1007/978-1-4939-8844-0Springer Science+Business Media, LLC, part of Springer Nature 2018.
- Kent, Alyssa G, Dupont, Chris L, Yooseph, Shibu, Martiny, Adam C 2016. Global biogeography of *Prochlorococcus* genome diversity in the surface ocean. *The ISME Journal* **10**: 1856–1865. <http://dx.doi.org/10.1038/ismej.2015.265>
- Kettler, G. C., Martiny, A. C., Huang, K., Zucker, J., Coleman, M. L., Rodrigue, S., Chen, F., Lapidus, A., Ferreira, S., Johnson, J., Steglich, C., Church, G. M., Richardson, P., Chisholm, S. W. 2007. Patterns and Implications of Gene Gain and Loss in the Evolution of *Prochlorococcus*. *PLoS Genetics* **3**: e231. <http://dx.doi.org/10.1371/journal.pgen.0030231>
- Larkin, A. A., Blinbry, S. K., Howes, C., Lin, Y., Loftus, S. E., Schmaus, C. A., Zinser, E. R., Johnson, Z. I. 2016. Niche partitioning and biogeography of high light adapted *Prochlorococcus* across taxonomic ranks in the North Pacific. *The ISME Journal* **10**: 1555–1567. <http://dx.doi.org/10.1038/ismej.2015.244>
- Larkin, A. A., C. A. Garcia, K. A. Ingoglia, N. S. Garcia, S. E. Baer, B. S. Twining, M. W. Lomas, and A. C. Martiny. 2019. Subtle biogeochemical regimes in the Indian Ocean revealed by spatial and diel frequency of *Prochlorococcus* haplotypes. *Limnol. Oceanogr.* **65**: S220–S232. doi:10.1002/lno.11251
- Lavin, P., P. Gomez, B. Gonzalez, and O. Ulloa. 2008. Diversity of the marine cyanobacteria *Prochlorococcus* and *Synechococcus* assessed by terminal restriction fragment length polymorphisms of 16S-23S rRNA internal transcribed spacer sequences. *Rev. Chil. Hist. Nat.* **81**: 515–531.
- Letunic, I., and P. Bork. 2019. Interactive tree of life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **47**: W256–W259. doi:10.1093/nar/gkz239
- Li, B., and others. 2010. Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc. Natl. Acad. Sci. USA* **107**: 10430–10435. doi:10.1073/pnas.0913677107
- Lynch, M. D. J., and J. D. Neufeld. 2015. Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* **13**: 217–229. doi:10.1038/nrmicro3400
- Malmstrom, R. R., A. Coe, G. C. Kettler, A. C. Martiny, J. Frias-Lopez, E. R. Zinser, and S. W. Chisholm. 2010. Temporal dynamics of *Prochlorococcus* ecotypes in the Atlantic and Pacific Oceans. *ISME J.* **4**: 1252–1264.
- Martiny, A., S. Kathuria, and P. Berube. 2009a. Widespread metabolic potential for nitrite and nitrate assimilation among *Prochlorococcus* ecotypes. *Proc. Natl. Acad. Sci. USA* **106**: 10787–10792.
- Martiny, A., A. Tai, D. Veneziano, F. Primeau, and S. W. Chisholm. 2009b. Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. *Environ. Microbiol.* **11**: 823–832. doi:10.1111/j.1462-2920.2008.01803.x
- Melo, D., G. Garcia, A. Hubbe, A. P. Assis, and G. Marroig. 2015. EvolQG - an R package for evolutionary quantitative genetics. *F1000Research* **4**: 925. doi:10.12688/f1000research.7082.3
- Moldovan, M. A., Gelfand, M. S. 2018. Pangenomic Definition of Prokaryotic Species and the Phylogenetic Structure of *Prochlorococcus* spp. *Front. Microbiol.* **9**: <http://dx.doi.org/10.3389/fmicb.2018.00428>
- Moore, L. R., G. Rocap, and S. W. Chisholm. 1998. Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393**: 464–467.
- Muñoz-Marín, M. C., G. Gómez-Baena, A. López-Lozano, J. A. Moreno-Cabezuelo, J. Díez, and J. M. García-Fernández. 2020. Mixotrophy in marine picocyanobacteria: Use of organic compounds by *Prochlorococcus* and *Synechococcus*. *ISME J.* **14**: 1065–1073. doi:10.1038/s41396-020-0603-9
- Oksanen, E. J., and others. 2019. Vegan: Community ecology package. R Package Version. R Package Version 2.5-6.
- Partensky, F., W. R. Hess, and D. Vaultot. 1999. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* **63**: 106–127.
- Paver, S. F., R. J. Newton, and M. L. Coleman. 2020. Microbial communities of the Laurentian Great Lakes reflect connectivity and local biogeochemistry. *Environ. Microbiol.* **22**: 433–446. doi:10.1111/1462-2920.14862
- Pittera, J., F. Humily, M. Thorel, D. Grulois, L. Garczarek, and C. Six. 2014. Connecting thermal physiology and latitudinal niche partitioning in marine *Synechococcus*. *ISME J.* **8**: 1221–1236. doi:10.1038/ismej.2013.228
- Ribalet, F., and others. 2015. Light-driven synchrony of *Prochlorococcus* growth and mortality in the subtropical Pacific gyre. *Proc. Natl. Acad. Sci. USA* **112**: 8008–8012.

- Rocap, G., D. Distel, J. B. Waterbury, and S. W. Chisholm. 2002. Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal internal transcribed spacer sequences. *Appl. Environ. Microbiol.* **68**: 1180–1191.
- Rocap, G., and others. 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**: 1042–1047.
- Roden, G. I. 1991. Subarctic-subtropical transition zone of the North Pacific: Large-scale aspects and mesoscale structure, p. 1–38. NOAA Technical Reports NMSF 105.
- Rusch, D. B., A. C. Martiny, C. L. Dupont, A. L. Halpern, and J. Craig Venter. 2010. Characterization of *Prochlorococcus* clades from iron-depleted oceanic regions. *Proc. Natl. Acad. Sci. USA* **107**: 16184–16189. doi:10.1073/pnas.1009513107
- Sher, D., J. W. Thompson, N. Kashtan, L. Croal, and S. W. Chisholm. 2011. Response of *Prochlorococcus* ecotypes to co-culture with diverse marine bacteria. *ISME J.* **5**: 1125–1132.
- Sohm, J. A., Ahlgren, N. A., Thomson, Z. J., Williams, C., Moffett, J. W., Saito, M. A., Webb, E. A., Rocap, G. 2016. Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. *ISME J.* **10**: 333–345. <http://dx.doi.org/10.1038/ismej.2015.115>
- Stecher, G., K. Tamura, and S. Kumar. 2020. Molecular evolutionary genetics analysis (MEGA) for macOS. *Mol. Biol. Evol.* **37**: 1237–1239. doi:10.1093/molbev/msz312
- Sullivan, M. B., J. B. Waterbury, and S. W. Chisholm. 2003. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* **424**: 1047–1051.
- Tai, V., R. S. Burton, and B. Palenik. 2011. Temporal and spatial distributions of marine *Synechococcus* in the Southern California Bight assessed by hybridization to bead-arrays. *Mar. Ecol. Prog. Ser.* **426**: 133–147.
- Thompson, A. 2020. Counts of picocyanobacteria from SMILE Cruise (SKQ201703S) at the North Pacific Subtropical Front in March 2017. Biological and Chemical Oceanography Data Management Office (BCO-DMO).
- Thompson, A. W., G. van den Engh, N. A. Ahlgren, K. Kouba, S. Ward, S. T. Wilson, and D. M. Karl. 2018. Dynamics of *Prochlorococcus* diversity and photoacclimation during short-term shifts in water column stratification at station ALOHA. *Front. Mar. Sci.* **5**: 488.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680. doi:10.1093/nar/22.22.4673
- Tschoeke, D., L. Vidal, M. Campeão, V. W. Salazar, J. Swings, F. Thompson, and C. Thompson. 2020. Unlocking the genomic taxonomy of the *Prochlorococcus* collective. *bioRxiv*, January, 2020.03.09.980698. doi:10.1101/2020.03.09.980698
- Whittaker, K. A., and T. A. Rynearson. 2017. Evidence for environmental and ecological selection in a microbe with no geographic limits to gene flow. *Proc. Natl. Acad. Sci. USA* **114**: 2651–2656. doi:10.1073/PNAS.1612346114
- Zinser, E. R., A. Coe, Z. I. Johnson, A. C. Martiny, N. J. Fuller, D. J. Scanlan, and S. W. Chisholm. 2006. *Prochlorococcus* ecotype abundances in the North Atlantic Ocean as revealed by an improved quantitative PCR method. *Appl. Environ. Microbiol.* **72**: 723–732. doi:10.1128/AEM.72.1.723-732.2006
- Zinser, E. R., Z. I. Johnson, A. Coe, D. Veneziano, and S. W. Chisholm. 2007. Influence of light and temperature on *Prochlorococcus* ecotype distributions in the Atlantic Ocean. *Limnol. Oceanogr.* **52**: 2205–2220.

Acknowledgments

We thank James Girton, Eric Kunze, and John Mickett, co-chief scientists of SKQ201703S (SMILE) in March–April 2017 (NSF OCE-1536314), for the invitation to participate and their analysis of physical parameters of the study region. We also thank the officers, crew, and scientists onboard the R/V *Sikuliaq* with special thanks to Bernard Mckiernan. Funding was provided by NSF OCE-1646709 to Thompson.

Conflict of Interest

None declared.

Submitted 02 July 2020

Revised 19 October 2020

Accepted 31 October 2020

Associate editor: Susanne Menden-Deuer