Sharp transitions in phytoplankton communities across estuarine to open ocean waters of the tropical Pacific

Sarah J. Tucker [1,2,3,#] https://orcid.org/0000-0003-0541-0853

Yoshimi M. Rii[1,3] https://orcid.org/0000-0002-6486-8955

Kelle C. Freel[1] https://orcid.org/0000-0001-8350-5327

Keli‘iahonui Kotubetey[4]

A. Hi‘ilei Kawelo[4]

Kawika B. Winter[1,3] https://orcid.org/0000-0003-3762-7125

Michael S. Rappé[1] https://orcid.org/0000-0002-9829-251X

[1]Hawai‘i Institute of Marine Biology, University of Hawai‘i at Mānoa, Kāne‘ohe, Hawai‘i, USA, 96744

[2]Marine Biology Graduate Program, University of Hawai‘i at Mānoa, Honolulu, Hawai‘i, USA, 96822

[3]He‘eia National Estuarine Research Reserve, Kāne‘ohe, Hawai‘i, USA, 96744

[4]Paepae o He‘eia, Kāne‘ohe Hawai‘i, USA, 96744

[#]Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA 02543, USA

***Corresponding Author**: Email: rappe@hawaii.edu

Running Title: Spatiotemporal shifts in phytoplankton

Abstract

Islands in the tropical Pacific supply elevated nutrients to nearshore waters that enhance phytoplankton biomass and create hotspots of productivity in otherwise nutrient-poor oceans. Despite the importance of these hotspots in supporting nearshore food webs, the fine-scale spatial and temporal variability of phytoplankton enhancement and changes in the underlying phytoplankton communities across nearshore to open ocean systems remain poorly understood. In this study, a combination of flow cytometry, pigment analyses, 16S rRNA gene amplicons, and metagenomic sequencing provide a synoptic view of phytoplankton dynamics over a four-year, near-monthly time-series across coastal Kāne‘ohe Bay, Hawai‘i, spanning from an estuarine Indigenous aquaculture system to the adjacent offshore environment. Through comparisons with measurements taken at Station ALOHA located in the oligotrophic North Pacific Subtropical Gyre, we elucidated a sharp and persistent transition between picocyanobacterial communities, from *Synechococcus* abundant in the nearshore to *Prochlorococcus* proliferating in offshore and open ocean waters. In comparison to immediately adjacent offshore waters and the surrounding open ocean, phytoplankton biomass within Kān‘eohe Bay was dramatically elevated. While phytoplankton community composition revealed strong seasonal patterns, phytoplankton biomass positively correlated with wind speeds, rainfall, and wind direction, and not water temperatures. These findings reveal sharp transitions in ocean biogeochemistry and phytoplankton dynamics across estuarine to open ocean waters in the tropical Pacific and provide a foundation for quantifying deviations from baseline conditions due to ongoing climate change.

Introduction

In marine ecosystems, phytoplankton play a crucial role by forming the base of the aquatic food web, where their productivity, abundance, cell size, and community composition are greatly influenced by light and nutrient availability (Azam et al. 1983). Surrounded by oligotrophic, open ocean waters, the coastal waters of remote islands in the tropical Pacific harbor a sharp increase in nutrients through physical oceanographic, biological, geological, and anthropogenic processes that results in increased phytoplankton biomass, cell size, and productivity (i.e. the Island Mass Effect, or IME; Doty and Oguri 1956; Gove et al. 2016). The enhanced primary productivity in turn promotes secondary productivity, supporting regional fisheries (Stock et al. 2017), increased biodiversity (Messié et al. 2022), and other marine resources relied upon by island communities.

Given the importance of elevated phytoplankton biomass from near-island coastal waters to maintaining healthy and productive coastal food webs, understanding the fine-scale variability of phytoplankton communities across coastal to open ocean systems adjacent to island masses can inform both the management of local marine environments and larger ecosystem models. Currently, phytoplankton biomass and productivity in the open oceans are rapidly changing. One result of increasing sea surface temperatures due to ongoing global climate change is an increase in the intensity of water column stratification of the open ocean, which can trap nutrients at depths below where phytoplankton at the ocean's surface can access them (Li et al. 2020). This has led to an expansion of nutrient-poor "ocean deserts" in the open ocean gyres and a decline in global phytoplankton biomass and primary productivity (Kwiatkowski et al. 2018). Declines in primary productivity will likely be amplified across the trophic food web in the near future, with an expected 14% decline in zooplankton biomass as soon as 2100 (Kwiatkowski et al. 2018) and a 20% decline in global fisheries by 2300 (Moore et al. 2018).

Small phytoplankton such as the two most abundant phytoplankton globally, *Prochlorococcus* and *Synechococccus*, are expected to increase in abundance under ocean warming conditions and decreased nutrient availability at the expense of larger sized phytoplankton groups (Flombaum et al. 2020). Small phytoplankton are often too small to be effectively grazed by metazoans, and are consumed by intermediary microzooplankton grazers who are then fed upon by larger zooplankton (Calbet and Landry 1999). Systems dominated by small phytoplankton have longer food chains and potentially a reduced energy transfer efficiency to higher trophic levels (Eddy et al. 2021). In contrast, large phytoplankton like diatoms can be consumed directly by zooplankton grazers (Calbet and Landry 1999) such as copepods, so that forage fish are only one trophic level apart from phytoplankton. Thus, changes in phytoplankton size structure and community composition also have profound implications for food web dynamics.

Unfortunately, the effects of increased open ocean stratification on food webs and biological productivity in adjacent coastal environments is uncertain. In part, this is due to studies predominantly focusing on coastal or oceanic systems in isolation (Xenopoulos et al. 2017), but also because the satellite-based methods that have led to an extensive understanding of primary productivity in the global open ocean have not yet been developed for shallow coastal waters (Carswell et al. 2017). Importantly, coastal marine food webs of islands situated in oligotrophic waters may be particularly vulnerable to the impacts of open ocean stratification.

Defining how nutrient availability and phytoplankton community composition and biomass vary with space and time in near-island and adjacent open ocean environments can provide a foundation for quantifying deviations from baseline conditions and predicting food web shifts because of climate change. To illuminate the factors influencing phytoplankton

communities across near-island to open ocean environments in the tropical Pacific, this study examined the effect of spatial and temporal variability in biogeochemical conditions on phytoplankton communities across multiple habitats that link the coastal environment of O‘ahu, Hawai‘i, with the offshore. These habitats span a tidally-influenced, estuarine environment within an Indigenous aquaculture system, through the interior of coastal Kāne‘ohe Bay, and to the offshore ocean environment surrounding Kāne‘ohe Bay. We also made comparisons to data collected by the Hawaii Ocean Time-series (HOT), a 30+ year time-series initiative measuring temporal trends of the adjacent ultraoligotrophic North Pacific Subtropical Gyre (NPSG; Karl and Church 2014). Together, this extensive spatial and temporal coverage revealed dramatic nearshore enhancement of phytoplankton biomass, pronounced seasonality in nearshore biogeochemistry and phytoplankton biomass and composition, and distinct transitions in phytoplankton communities spanning <6 km to >100 km across Kāne‘ohe Bay to the NPSG.

Materials and Methods

Study location

The Hawaiian archipelago within the oligotrophic NPSG is the world’s most remote island chain. Kāne‘ohe Bay, located on the windward side of the island of O‘ahu (21° 28′ N, 157° 48′ W), is a well-studied, coral-reef-dominated embayment (**Fig. 1a**). The bay has a total surface area of 41.4 km^2 and is approximately 4.3 km wide, 12.8 km in length, and 10 m deep on average (Jokiel 1991). Sharp nearshore to offshore gradients in biogeochemical parameters occur over a short distance (<6 km) along with a diverse topography due to patch, fringing, and barrier reefs (Jokiel 1991; Tucker et al. 2021). Localized freshwater input from streams contribute to episodic spatial variability in environmental conditions, including salinity and inorganic nutrient concentrations (Cox et al. 2006; Yeo et al. 2013; Tucker et al. 2021). Water residence time

111 within the bay varies from less than a day to over one month (Lowe et al. 2009), with the highest
112 residence times in the sheltered southern lobe. Oceanic water, primarily driven by wave action,
113 flows into the bay over a large barrier reef located in the central bay (Lowe et al. 2009). Water is
114 generally transported out of the bay through two nearly-parallel channels positioned in the
115 southern and northern portions of the bay. For most of the year, the bay is well mixed by
116 tradewinds. However, periods of high temperatures and low wind speeds can cause vertical
117 stratification in the water column (Smith 1981).
118
119 **Collaboratively developed research**
120 At the mouth of He‘eia Stream in the southern section of Kāne‘ohe Bay is an ~800-year-
121 old, 0.356 km^2 Indigenous aquaculture system known contemporarily as He‘eia Fishpond, but
122 anciently as Pihi Loko I‘a (Kelly 1973). Indigenous aquaculture systems in Hawai‘i engaged in
123 trophic engineering to promote primary productivity that sustained the population through
124 abundant food fish and reef fish (Winter et al. 2020a). A 2.5 km basalt rock wall filled with coral
125 rubble encompasses He‘eia Fishpond. The wall is equipped with multiple sluice gates that
126 increase water residence time while still allowing for exchange between coastal Kāne‘ohe Bay
127 and He‘eia stream waters. The sluice gate (mākāhā) system allows juvenile fish to flourish
128 within the high nutrient environment of the fishpond, while protecting them from predators.
129 While it has been understood that the fishpond provided the perfect nursery habitat for prized
130 food fish, little is known about the exchange of organisms from within the fishpond to the rest of
131 Kāne‘ohe Bay, and beyond. With this interest expressed by the Indigenous stewards of He‘eia
132 Fishpond, we sought to establish a current, baseline understanding of phytoplankton biomass and
133 community composition from within the estuarine fishpond environment out to open ocean

waters adjacent to O‘ahu. The methods employed in this study were collaboratively developed with He‘eia Fishpond stewards and the He‘eia National Estuarine Research Reserve (NERR; Winter et al. 2020b). All individuals who collaboratively developed the methods also collaborated on interpreting the data and are listed as co-authors on this paper. Sampling campaigns were conducted with permission from Paepae o He‘eia, the stewards of He‘eia Fishpond, and the private landowner, Kamehameha Schools.

Sample collection and environmental parameters

Between August 2017 and June 2021, seawater was collected from a depth of 2 m at 10 sites in Kāne‘ohe Bay and the adjacent offshore waters on a near-monthly basis (36 sampling events over 46 months) as part of the Kāne‘ohe Bay Time-series (KByT) using previously described methods (**Fig. 1a, Supporting Information Table S1**; Tucker et al. 2021). Between September 2020 and June 2021, two additional stations within He‘eia Fishpond were also sampled at a quarterly interval **(Fig. 1a**): Station Wai2 (now known as Waimā‘ama), located in the northwestern corner of the Fishpond within the He‘eia Stream mouth (known as muliwai) is a highly turbid and brackish water environment with tidal fluctuations resulting in salinity ranges of 10-35 ppt. Station Kaho‘okele (Kaho) is located at a sluice gate facing the ocean and receives high exchange with coastal Kāne‘ohe Bay (~30-35 ppt, Möhlenkamp et al. 2019). At all stations, seawater samples for biogeochemical analyses and nucleic acids were collected, as were *in situ* measurements of seawater temperature, pH, and salinity with a YSI 6600 or ProDSS multi-parameter sonde (YSI Incorporated, Yellow Springs, OH, USA). Approximately one liter of seawater was prefiltered with 85-µm Nitex mesh and subsequently filtered through a 25-mm diameter, 0.1-µm pore-sized polyethersulfone (PES) filter membrane (Supor-100, Pall Gelman

158 Inc., Ann Arbor, MI, USA) to collect microbial cells for DNA isolation. The filters were
159 subsequently submerged in DNA lysis buffer (Suzuki et al. 2001; Yeo et al. 2013) and stored in
160 −80°C until further processing.

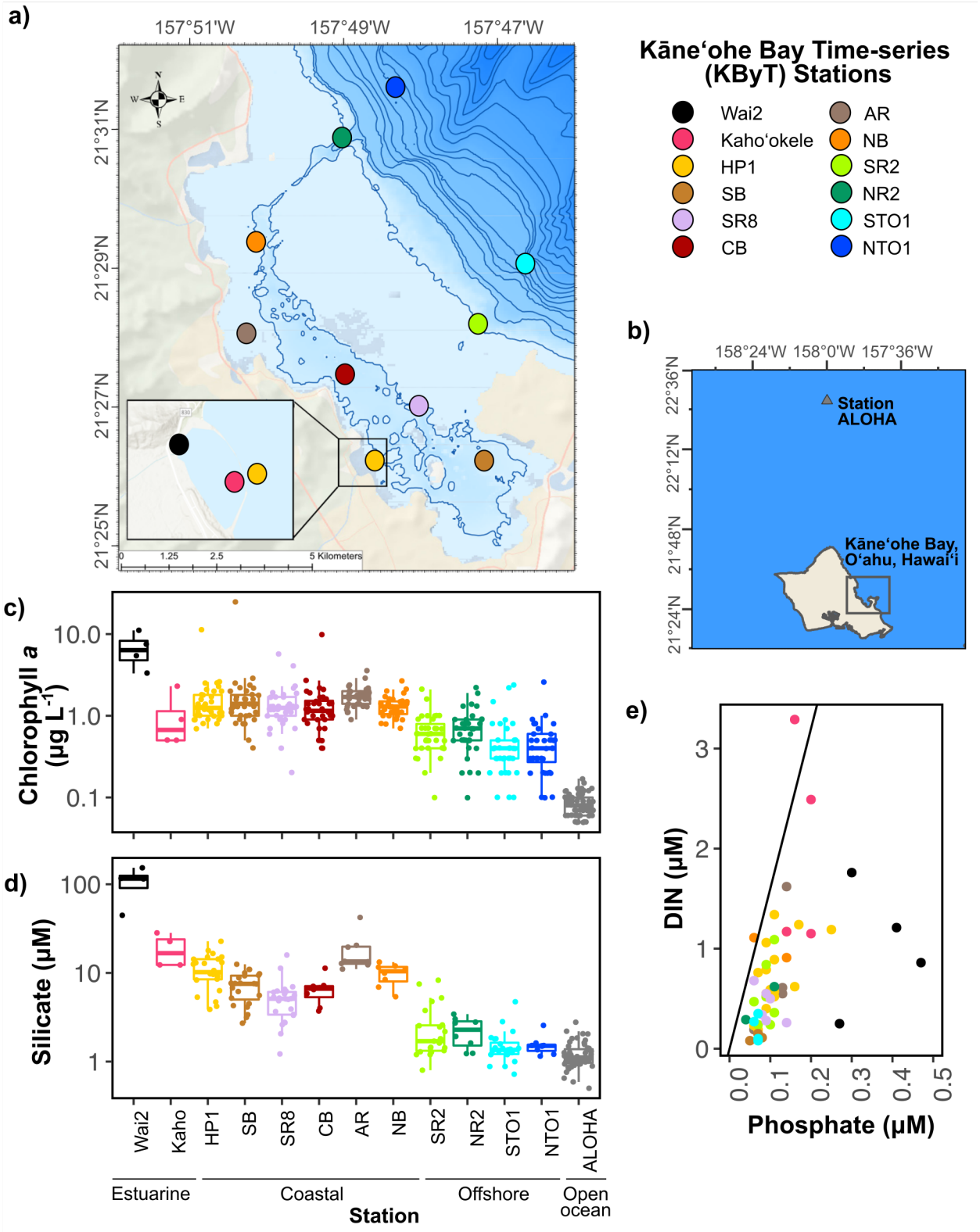


Fig. 1. a) Map of sampling stations located on the windward side of the island of Oʻahu Hawaiʻi. Inset shows two stations within Heʻeia Fishpond and one immediately adjacent to the fishpond.

164 Contour lines mark every ten meters up until 50 m and then 50 m intervals for depths >50 m. b)
165 Location of Kāneʻohe Bay on Oʻahu, Hawaiʻi and the position of Station ALOHA, a sampling
166 station of the Hawaii Ocean Time-series (HOT) program. c) Chlorophyll *a* and d) silicate
167 concentrations (both plotted on a log scale) from the estuarine to open ocean environments
168 examined in this study. e) Ratios of dissolved inorganic nitrogen (DIN):Phosphate. Diagonal line
169 denotes Redfield ratio of 16:1 suggesting nitrogen limitation is characteristic of the system.
170
171 Seawater subsamples for fluorometric chlorophyll *a* concentrations (125 mL) and
172 photosynthetic pigments via high-performance liquid chromatography (HPLC; 2 L) were
173 collected on 25-mm diameter GF/F glass microfiber filters (Whatman, GE Healthcare Life
174 Sciences, Chicago, IL, USA) and stored in aluminum foil at −80°C until extraction. The
175 collection of phytoplankton pigments on the GF/F glass microfiber filters allow for comparisons
176 with the Hawaii Ocean Time-series data. However, because the filters have a pore size of 0.7µm,
177 we acknowledge that most small cyanobacteria were likely missed. Chlorophyll *a* was extracted
178 with 100% acetone and measured with a Turner 10-AU fluorometer (Turner Designs, Sunnyvale,
179 CA, USA) following standard techniques (Welschmeyer 1994). Photosynthetic pigments
180 measured via high performance liquid chromatography were extracted in 100% acetone and
181 analyzed on a Waters 2690 separations module equipped with a C18 column and full spectrum
182 photodiode array detector, following (Mantoura and Llewellyn 1983) and modified according to
183 (Bidigare et al. 1989). Chlorophyll *a* concentrations measured via the fluorometer are herein
184 referred to chlorophyll *a* (Chla), while chlorophyll *a* concentrations measured by high
185 performance liquid chromatography are specified as total chlorophyll *a* (TChla).

For cellular enumeration, seawater was preserved in 2 mL aliquots in a final concentration of 0.95% (v:v) paraformaldehyde (Electron Microscopy Services, Hatfield, PA, USA) at −80 °C until analyzed via flow cytometry. Cellular enumeration of cyanobacterial picophytoplankton (*Synechococcus* and *Prochlorococcus*), eukaryotic picophytoplankton, and non-cyanobacterial (presumably heterotrophic) bacteria and archaea (hereafter referred to as heterotrophic bacteria) was performed on a Beckman Coulter CytoFLEX S, following the method of (Monger and Landry 1993). Inorganic nutrients were measured using a Seal Analytical AA3 HR Nutrient Autoanalyzer (detection limits: $NO_2^- + NO_3^-$, 0.009 µM; SiO_4, 0.09 µM; PO_4^{3-}, 0.009 µM; NH_4, 0.03 µM).

Rainfall, wind speed, and wind direction were monitored using data collected at a meteorological station located at the Hawai‘i Institute of Marine Biology (HIMB) on Moku o Lo‘e in Kāne‘ohe Bay (http://www.pacioos.hawaii.edu/weather/obs-mokuoloe/). Either maximum (e.g., rainfall, wind speed) or average (wind direction) values were taken across 1- to 7-day windows leading up to the sampling event, depending on data availability from the station. One sampling event (February 5, 2021) had no data for the 7 days prior to sampling and so the data for a 30-day window were used. Metadata from Station ALOHA (22° 45’N, 158° 00’W, ~5 m depth, August 2017 to December 2020), a sampling station of the HOT program (Karl and Church 2014), were downloaded from https://hahana.soest.hawaii.edu/hot/hot-dogs/ (accessed on 9/12/2022).

Spatiotemporal comparisons of environmental variables, cellular abundances, and phytoplankton pigments were conducted using the R package ‘multcomp’ (Hothorn et al. 2008) with one-way ANOVAs testing for multiple comparisons of means with Holm correction and Tukey contrasts. Summer (28 June through 28 September) and winter (27 December through 29

March) seasons were defined using harmonic regression analyses of surface seawater temperature collected hourly between 2010–2019 at NOAA station MOKH1 in Kāneʻohe Bay (https://www.ndbc.noaa.gov/station_page.php?station=mokh1; Tucker et al. 2021). Spatiotemporal variation in cellular abundances and phytoplankton pigments were visualized using 'mba.surf' from MBA (Finley et al. 2017) to interpolate data over the KByT sampling events and stations.

The map of Kāneʻohe Bay was plotted in ArcGIS Pro v2.9. Contour lines were drawn using bathymetry metadata (http://www.soest.hawaii.edu/hmrg/multibeam/bathymetry.php). The map of Oʻahu was plotted in R with 'geom_sf' from ggplot2 (Wickham 2016), using shape file from the Hawaii Statewide GIS Program (https://prod-histategis.opendata.arcgis.com/maps/HiStateGIS::coastline).

DNA extraction, 16S rRNA gene amplicon sequencing, & metagenome sequencing

DNA extraction and 16S rRNA gene sequencing followed previously published methods (Tucker et al. 2021). Briefly, amplicon libraries were made from polymerase chain reactions of the 16S rRNA gene using barcoded 515F and 926R universal primers (Parada et al. 2016) and paired-end sequenced with MiSeq v2 2x250 technology (Illumina, San Diego, CA, USA). Genomic DNA from a subset of 32 of the 368 total samples collected between 2017-2021 were used for metagenomic sequencing. This included samples from four sampling events between 2017 and 2019 at 6-10 stations. Libraries were constructed from approximately 100 ng of genomic DNA using the Kappa HyperPrep Kit (Roche, Pleasanton, CA, USA) with mechanical shearing (Covaris, Woburn, MA, USA) and paired-end sequenced on a single lane of the NovaSeq 6000 SP 150 (Illumina, San Diego, CA, USA).

Sequence analysis

Amplicon sequence data generated from KByT sampling between July 2019 and June 2021 were analyzed in conjunction with previously published amplicon data spanning August 2017 to June 2019 (PRJNA706753; (Tucker et al. 2021). For each of the two sequencing runs, samples were demultiplexed and quality controlled using Qiime2 v2.4 (Bolyen et al. 2019). Full length forward reads (251 base pairs) were denoised using DADA2 (Callahan et al. 2019) to delineate amplicon sequencing variants (ASVs). Reverse reads were not used because of inconsistent quality. ASVs were assigned taxonomy using SILVA v138 as a reference database (Quast et al. 2012) and the two runs were subsequently merged in Qiime2. ASVs that contained at least 10 reads in at least two samples were retained.

ASVs classified by the SILVA v138 database as Eukaryota, unassigned at the domain level, or classified as chloroplast at the order level were re-classified using the PR2 v 4.14.0 database (Guillou et al. 2013) in the DECIPHER R package (Wright 2016) using a 60% confidence threshold cut off. Sequences classified as Bacteria, Archaea, or chloroplast at the order-level were retained for further analyses, while those unclassified at the domain level or classified as Eukaryota were excluded from further analyses. In the context of amplicon sequence data, “phytoplankton” herein refers to ASVs classified as cyanobacteria and eukaryotic plastid sequences, although we recognize that mixotrophic and phagotrophic lifestyles may be included in this broad definition.

Statistical analyses were conducted using the R packages phyloseq (McMurdie and Holmes 2013), ggplot2 (Wickham 2016), pheatmap (Kolde 2019), and microbiome (Lahti and Shetty 2017). An Aitchison distance (Aitchison 1982), the Euclidean distance between centered log-ratio (clr)-transformed compositions, was used on the entire quality-controlled dataset of

255 phytoplankton raw read counts using 'transform' in the microbiome package (Lahti and Shetty
256 2017). Ward D2 hierarchical clustering using 'hclust ( )' in the stats base package of R was
257 applied to this matrix to cluster samples with amplicon data into groups and visualized with
258 dendextend (Galili 2015). DESeq2 (Love et al. 2014) was used to model differential abundance
259 patterns of amplicon data across environmental clusters using Wald Tests and Bonferroni
260 correction for multiple comparisons (alpha cutoff < 0.05). Divnet (Willis and Martin 2020) was
261 used to estimate differences in alpha diversity and test for significance between spatiotemporal
262 groupings. Pearson's correlation analyses were conducted in corrplot (Wei and Simko 2021).

263 Lomb Scargle Periodograms (LSP) in the lomb package (Ruf 1999) were used to define
264 seasonality among phytoplankton genera by determining the spectrum of frequencies in a
265 dataset: this approach can account for unevenly sampled time-series data and has been
266 previously applied to microbiome time-series analyses (Auladell et al. 2021). Only genera with
267 annual intervals (peak frequency = 1±0.25, $p<0.01$) as their most significant periodic trend were
268 considered as having seasonality. A starting frequency of 0.16 was used so as to not include
269 periodic components between two consecutive months. An inverse hyperbolic sine
270 transformation (asinh) was conducted on sequence data prior to LSP. Significance (q-
271 values<0.05) was corrected for multiple-testing using data randomization for LSP analyses using
272 fdrtools (Strimmer 2008).

273 **Metagenomic read recruitment**

274 To investigate the dominant cyanobacteria within and surrounding Kāne'ohe Bay, we
275 conducted metagenomic read recruitment of 32 metagenomes from KByT and 12 previously
276 published from the open ocean Station ALOHA (PRJNA352737; Mende et al. 2017) to 56
277 cyanobacterial genomes from *Prochlorococcus* (six minor clades) and the three major lineages of

the marine *Synecchococcus*/*Cyanobium* lineage [SC 5.1 (14 minor clades), SC 5.2, and SC 5.3] (**Supporting Information Table S2**). SC 5.2 is the only clade with both *Synechococcus* and *Cyanobium* members (Doré et al. 2020).

A contig database of the 56 cyanobacteria isolate genomes was constructed using anvi'o v 8.0 (Eren et al. 2021) following previously described pipelines (Delmont and Eren 2018). Briefly, Prodigal v2.6.3 (Hyatt et al. 2010) was used to identify open reading frames (ORFs) from the contigs and an anvi'o database was created using 'anvi-gen-contigs-db'. Metagenomic reads were first quality filtered using an Illumina-utils library v1.4.1 called 'iu-filterquality-minoche' (Eren et al. 2013) that uses quality filtering parameters described previously (Minoche et al. 2011). Quality filtered metagenomic reads were competitively mapped with Bowtie2 v2.3.5 (Langmead and Salzberg 2012) to an anvi'o contig database of cyanobacterial isolate genomes. The 'anvi-profile' function stored coverage and detection statistics of each cyanobacterial genomes found in the KByT and Station ALOHA metagenomic samples.

To evaluate the distribution of individual genomes, a "detection" metric, the proportion of the nucleotides in a given sequence that are covered by at least one short read, was used to evaluate if a population was present in a metagenomic sample. A detection value of at least 0.25 was used as a criterion to eliminate false positives, when an isolate genome was falsely found within a sample (Utter et al. 2020). Mean coverage Q2Q3, which refers to the average depth of coverage excluding nucleotide positions with coverages in the 1st and 4th quartiles, was mapped for each genome. Mean coverage Q2Q3 was summed across all cyanobacterial genomes per sample and then the genome (or all genomes in a clade) was divided by this sum to determine a relative abundance of a genome (or a clade) in each sample. Average nucleotide identity (ANI) was calculated using pyANI (Pritchard et al. 2015). A phylogenomic tree was estimated from the

56 cyanobacterial isolates and an outgroup using GTotree v1.4.16 (Lee 2019) with cyanobacterial single-copy genes and visualized in FigTree v 1.4.4 (http://tree.bio.ed.ac.uk/).

Data & code availability

Sequencing data are available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject number PRJNA706753 as well as PRJNA971314. Environmental data were submitted to BCO-DMO under https://www.bco-dmo.org/project/663665. Code used in the analysis is available at https://github.com/tucker4/Tucker_Phytoplankton_KByT_HeNERR.

Results

Biogeochemical parameters

Along the nearshore to open ocean waters of the tropical Pacific, biogeochemical parameters sharply declined across both small spatial scales and vast stretches of ocean (**Supporting Information Table S1)**. On average, chlorophyll *a* concentrations at stations in the coastal waters of Kāne‘ohe Bay increased 18-fold (1.6±1.9 *vs.* 0.09±0.03 µg L^{-1}, mean±sd) from the open ocean and 3-fold (1.6±1.9 *vs.* 0.6±0.4 µg L^{-1}) from the immediately adjacent offshore waters. The estuarine waters of He‘eia Fishpond harbored higher chlorophyll *a* concentrations compared to coastal stations (3.9±3.8 *vs.* 1.6±1.9 µg L^{-1}, **Fig. 1c**). Mean chlorophyll *a* concentrations increased 43-fold between the estuarine waters of He‘eia Fishpond and the open ocean (3.9±3.8 *vs.* 0.09±0.03 µg L^{-1}, **Supporting Information Table S1**) and 7-fold over the <6 km distance covering the interior and surrounding waters of Kāne‘ohe Bay (3.9±3.8 *vs.* 0.6±0.4 µg L^{-1}; **Fig. 1c**, **Supporting Information Table S1**). Stations offshore from Kāne‘ohe Bay had elevated concentrations of chlorophyll *a* compared to the open ocean (7-fold increase; 0.6±0.4 *vs.* 0.09±0.03 µg L^{-1}). Elevated phytoplankton biomass was a persistent feature within Kāne‘ohe

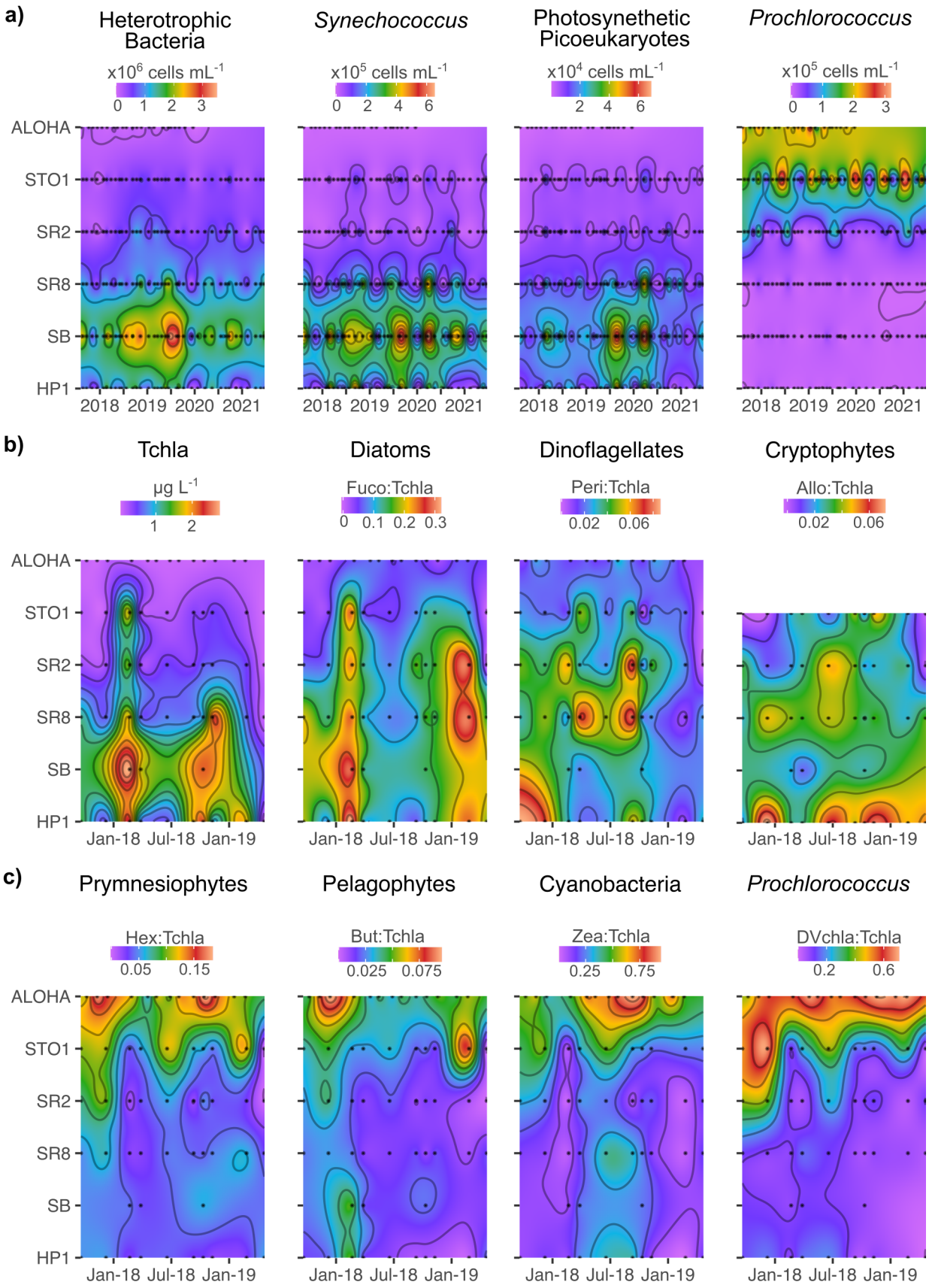
Bay, with increased chlorophyll *a* concentrations detected in at least one of the stations positioned in the coastal environment compared to the stations offshore during all 36 sampling events (**Supporting Information Fig. S1**).

Elevated concentrations of inorganic nutrients (silicate, nitrate+nitrite, phosphate, ammonia) were also found in the nearshore waters of Kāneʻohe Bay compared to offshore and open ocean stations (**Supporting Information Table S1 & S3**). Mean silicate concentrations at Wai2 of the estuarine stations were 107.3 µM, compared to 42.2 µM in the coastal stations. Across all KByT stations, phosphate and silicate concentrations differed significantly but were positively correlated with increasing chlorophyll *a* concentrations. However, nitrate+nitrite concentrations did not correlate with chlorophyll *a* concentrations (**Supporting Information Fig. S2**). Despite the overall increase of inorganic nutrients in the estuarine and coastal stations, all stations were below 16:1 N:P ratios using dissolved inorganic nitrogen (nitrate+nitrite plus ammonia) and phosphate (**Fig. 1e**, (Redfield 1960).

Microbial cell counts and phytoplankton pigments

In contrast to coastal stations where *Synechococcus* cellular abundance was high, *Prochlorococcus* cellular abundance was elevated in the stations positioned in the offshore waters surrounding Kāneʻohe Bay and at Station ALOHA (**Fig. 2a, Supporting Information Table S1 & S3**). Cellular abundances of heterotrophic bacteria and eukaryotic picophytoplankton were also greater in the coastal stations compared to offshore (**Fig. 2a, Supporting Information Table S1 & S3**). In coastal Kāneʻohe Bay, ratios of fucoxanthin, peridinin, and alloxanthin to total chlorophyll *a* (Tchla) concentrations were higher than in the offshore, indicating an increase in diatoms, dinoflagellates, and cryptophytes (respectively)

347 closer to shore (**Fig. 2b, Supporting Information Table S4**). In contrast, pigments relative to
348 Tchla for photosynthetic pigments diagnostic of prymnesiophytes (19′-hexanoyloxyfucoxanthin),
349 pelagophytes (19′-butanoyloxyfucoxanthin), cyanobacteria (zeaxanthin), and *Prochlorococcus*
350 (divinyl chlorophyll *a*) were higher in the offshore stations compared to the coastal environment
351 (**Fig. 2b**, **Supporting Information Table S4**).



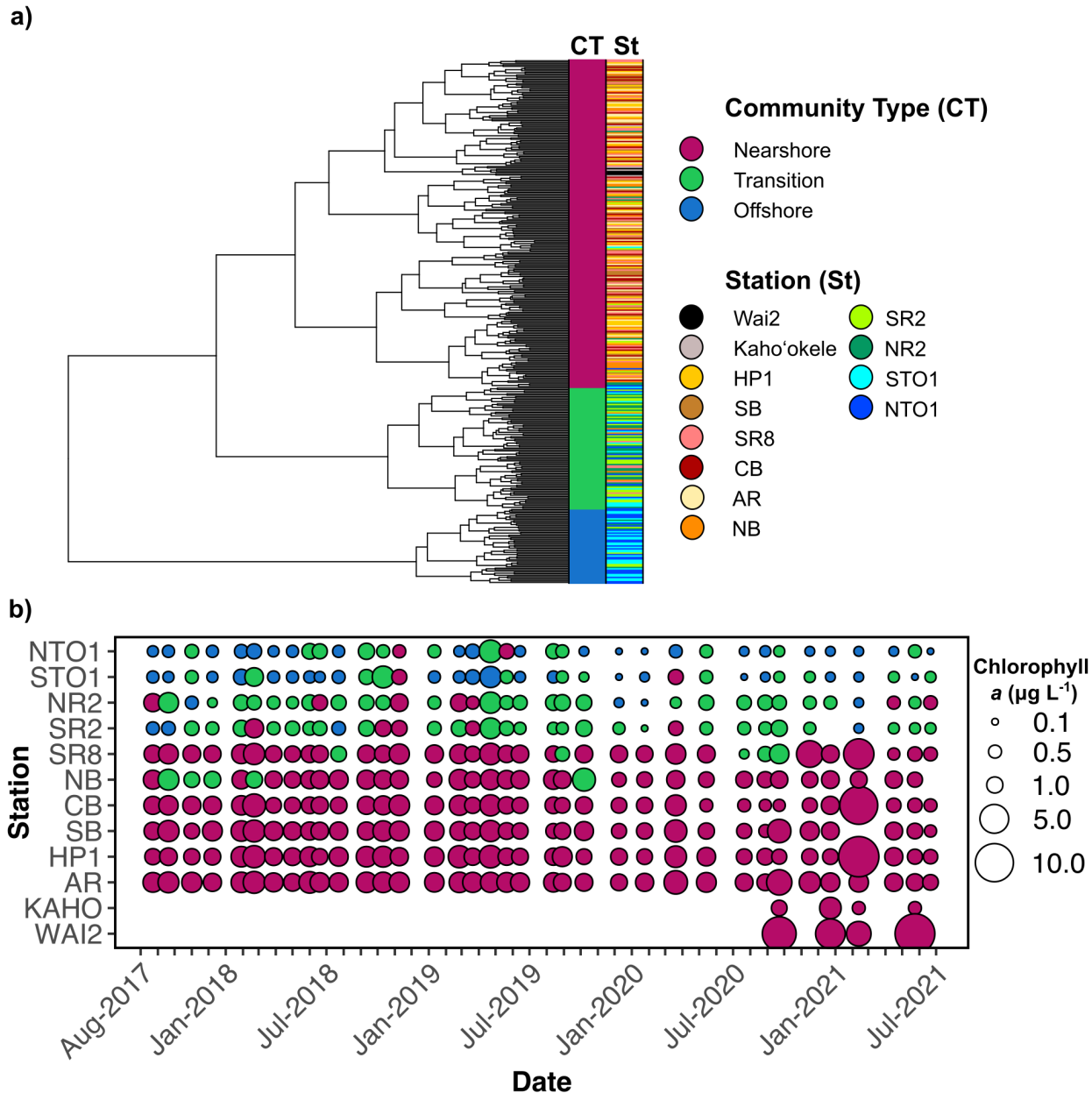
352

353 **Fig. 2.** Microbial cellular abundances and pigment concentration vary through time and space
354 across stations from the southern sector of Kāne‘ohe Bay (HP1, SB, SR8), offshore stations (SR2
355 and STO1), and open ocean Station ALOHA: a) Cellular abundances (cells mL^{-1}) of
356 heterotrophic bacteria, *Synechococcus*, photosynthetic picoeukaryotes, and *Prochlorococcus*; b)
357 Total chlorophyll *a* (Tchla) concentrations (µg L^{-1}) and ratios of phytoplankton pigments
358 indicative of specific phytoplankton groups relative to Tchla. Note, alloxanthin was below
359 detection levels for Station ALOHA and not presented here. Abbreviations: Fuco: Fucoxanthin,
360 Peri: Peridinin; Allo: Alloxanthin; Hex: 19′-hexanoyloxyfucoxanthin; But: 19′-
361 butanoyloxyfucoxanthin; Zea: zeaxanthin; Dvchla: divinyl chlorophyll *a*.
362

363 **Phytoplankton community composition through 16S rRNA gene sequencing**
364
365 We delineated 505 phytoplankton ASVs across 366 samples, including 66 from
366 cyanobacteria and 439 from eukaryotic plastids. Examining the distribution of phytoplankton
367 ASVs across samples revealed that phytoplankton communities clustered into three major
368 community types that coincide with spatial differences in biogeochemistry (**Fig. 3**, **Supporting**
369 **Information Table S5**). We categorized the three major community types as nearshore,
370 transition, and offshore because of their distinct biogeochemical characteristics and geographic
371 location (**Supporting Information Table S5, Fig. S3**). The nearshore cluster of 229 samples
372 included all samples collected from six stations found most closely located to land (Wai2,
373 Kaho‘okele, SB, CB, HP1, AR) and at least one sample from the six remaining stations. The
374 transition cluster of 85 samples consisted of samples collected from stations not immediately
375 next to land (NB, SR8, NR2, SR2, STO1, NTO1), while the offshore cluster encompassed 52

376 samples collected exclusively from the four stations located the furthest distance from land (SR2,

377 NR2, STO1, NTO1).



378

379 **Figure 3.** a) Hierarchical clustering of phytoplankton communities from 366 samples collected

380 at 12 sampling stations form three major groups, hereafter referred to as community types:

nearshore, transition, and offshore. b) The distribution of samples defined as nearshore, transition, and offshore community types across stations and sampling events (36 total between 2017-2021) shows spatial and temporal persistence in the distinct community types. The size of the circle represents the chlorophyll *a* concentration during the time of sampling and the color of the circle represents the community type: nearshore, transition, and offshore.

ASV richness was the highest in the nearshore, while the transition community group had the highest Shannon's diversity estimate (**Supporting Information Table S6**). Phytoplankton relative abundance was dominated by a few highly abundant groups, including *Synechococcus*, *Prochlorococcus*, Mamiellophyceae (green algae), and Bacillariophyta (diatoms) (**Fig. 4a)**. Using DESeq2 variance stabilized abundances, 20 classes (**Supporting Information Table S7**) and 33 genera of phytoplankton (or groups unclassified at genus-level but classified at the family-level; **Fig. 4b**, **Supporting Information Table S8**) differed significantly in abundance across the three community types.

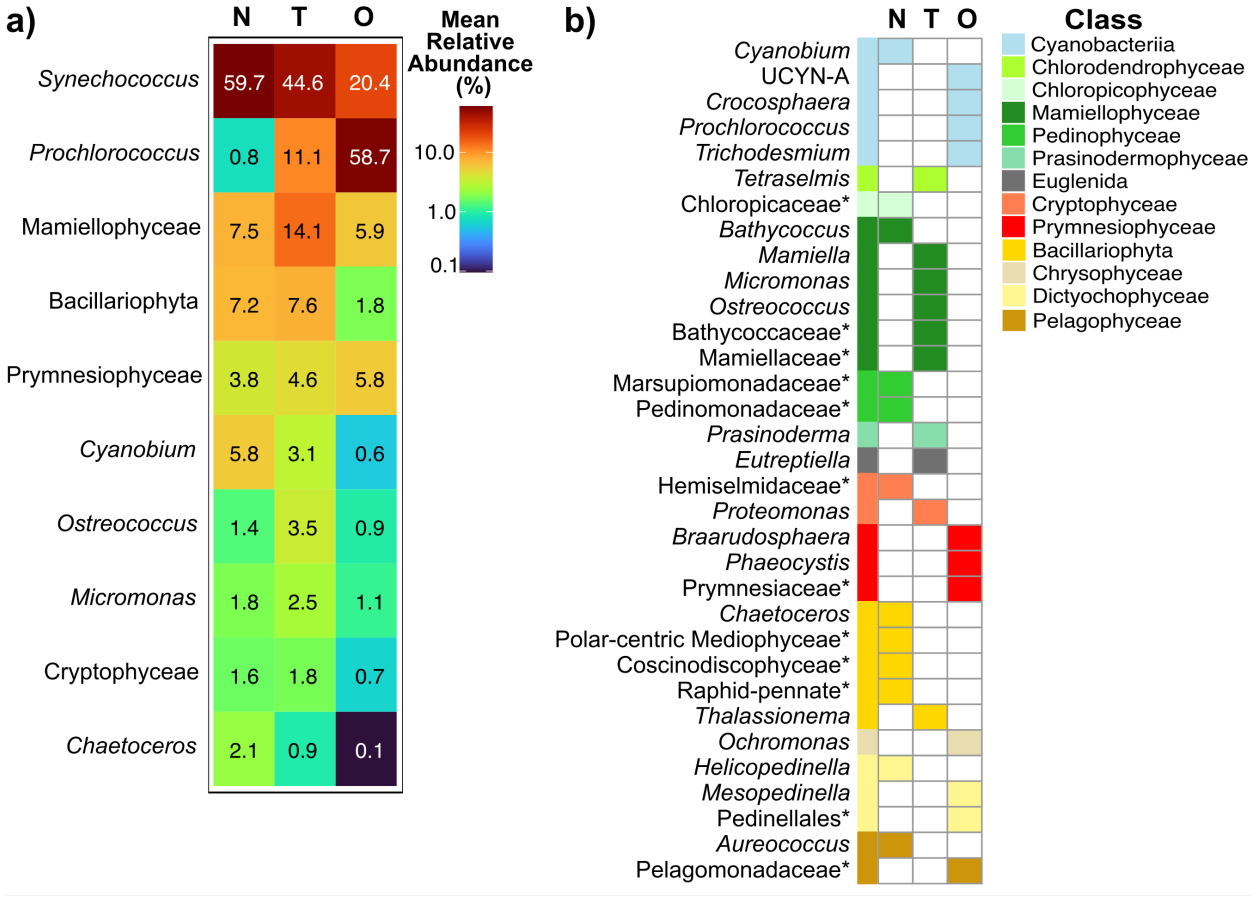


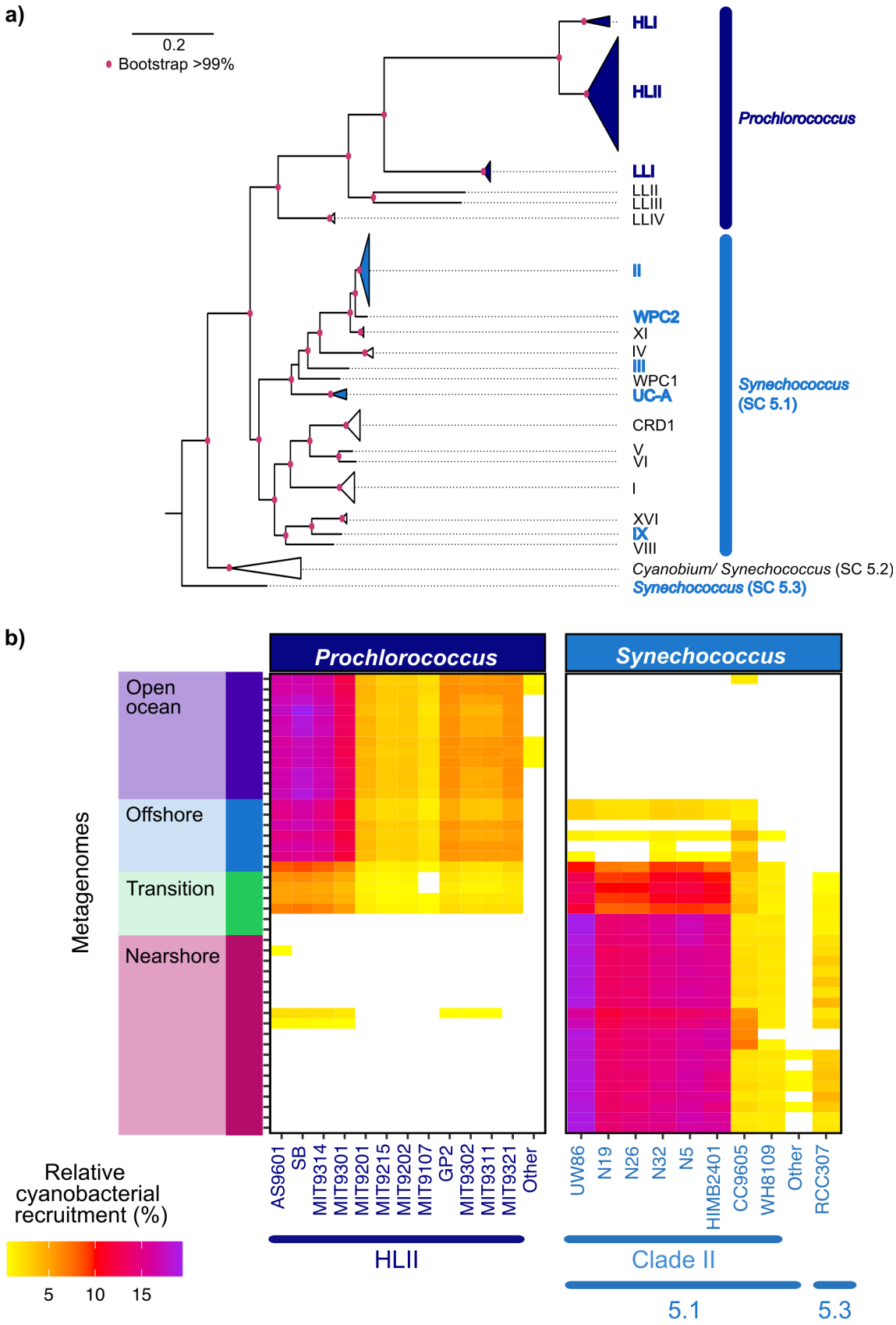
Fig. 4. a) Mean relative abundance of the top 10 phytoplankton groups across the nearshore (N), transition (T), and offshore (O) environments. b) Phytoplankton genera with significantly different distributions across the three environments. Colored boxes denote the peak in abundance for each genus. Phytoplankton groups that were classified at the family-level but unidentified at the genus-level are denoted with an asterisk.

Cyanobacterial population structure

Metagenomic read recruitment to 56 genomes of the cyanobacterial genera *Prochlorococcus*, *Synechococcus*, and *Cyanobium* showed that *Prochlorococcus* HLI, HLII, and LLI, and *Synechococcus* SC 5.1 II, WPC2, III, UC-A, and IX and SC 5.3 were detected in

metagenomes from the Kāne‘ohe Bay Time Series and surface ocean samples from Station ALOHA (**Fig. 5a**). Although *Cyanobium* 16S rRNA gene ASVs were detected in the amplicon data, no *Cyanobium* representatives (SC 5.2) were detected in our metagenomic read recruitment. Genomes from *Synechococcus* SC 5.1 II and SC 5.3 and *Prochlorococcus* HLII were among the most abundant representatives within our samples (**Fig. 5b**). *Prochlorococcus* HLII comprised 98.9±1.0% (mean±sd) of cyanobacterial relative abundance in the open ocean and 83.1±17.4% of the cyanobacterial relative abundance in the offshore. *Prochlorococcus* HLII recruited only a small proportion of reads from a handful of nearshore samples, where it made up 1.1±2.4% of the cyanobacterial relative abundance. *Synechococcus* clade II comprised 96.3±2.5% of the cyanobacterial relative abundance in the nearshore Kāne‘ohe Bay community type. *Synechococcus* SC 5.3 also recruited some metagenomic reads, but only from coastal Kāne‘ohe Bay samples and at low relative cyanobacterial abundance (2.4±1.2%) (**Fig. 5b**).

Within the *Prochlorococcus* HLII and *Synechococcus* II clades, read recruitment varied between closely related genomes. Read recruitment was substantially higher in *Synechococcus* clade II isolate UW86 compared to all other clade II genomes, despite sharing >95% ANI with most other clade II genomes (**Supporting Information Fig. S4**). Within *Prochlorococcus* HLII, isolate genomes AS9601, SB, MIT9314, and to a lesser extent MIT9301, recruited a substantially greater proportion of reads than other members of this clade. AS9601, SB, MIT9314, and MIT9301 shared 94% ANI, which is higher than what was shared with other HLII isolates (<93%), with the exclusion of MIT9215 and MIT9202 who share 97% ANI with each other (**Supporting Information Fig. S4**).



430

Fig. 5. a) Phylogenomic tree based on cyanobacterial marker genes found in 56 *Prochlorococcus*, *Synechococcus*, and *Cyanobium* isolate genomes and outgroup (*Gloeobacter violaceus*- not shown). Clades detected in surface ocean metagenomic samples from the Kāne‘ohe Bay Time-series (KByT) and Station ALOHA are colored and bolded. b) The relative abundance of cyanobacterial genomes across KByT and Station ALOHA metagenomes was dominated by *Prochlorococcus* HLII, *Synechococcus* II (SC 5.1), and *Synechococcus* SC 5.3. Recruitment of <0.5% not shown.

Seasonality in biogeochemistry and community composition

Surface seawater temperatures were significantly cooler in the winter than summer in nearshore Kāne‘ohe Bay and the adjacent transition and offshore waters (**Fig. 6a, Supporting Information Table S9**). Cellular concentrations of *Synechococcus* and heterotrophic bacteria were higher in the summer than the winter in the nearshore, transition, and offshore (**Fig.6a, Supporting Information Table S9**). Photosynthetic picoeukaryote cellular concentrations increased during the summer in coastal and transition environments (coastal: $p=0.004$, transition: $p<0.001$, **Fig. 6a, Supporting Information Table S9**), but did not vary seasonally in the offshore. *Prochlorococcus* cell concentrations only varied seasonally in the nearshore where it increased in the winter ($p<0.001$, **Fig. 6a, Supporting Information Table S9**).

Chlorophyll *a* concentrations increased in the winter compared to the summer in the nearshore (mean±sd; winter: 2.3±3.4 μg L^{-1}, summer: 0.8±0.5 μg L^{-1}; $p=0.031$; **Fig. 6a, Supporting Information Table S9**), but not in the transition or offshore waters. Importantly, given that the collection method for chlorophyll *a* concentrations likely missed most small cyanobacteria, it is possible that seasonality has been underestimated. Nearshore chlorophyll *a*

concentrations increased with wind speed, wind direction, and rainfall, but not with seawater temperature (**Fig. 6b**). Three sampling events with elevated chlorophyll *a* concentrations in the nearshore occurred during anomalously high wind speeds or rainfall events (**Supporting Information Fig. S5**). Chlorophyll *a* concentrations in the transition and offshore did not correlate with seawater temperatures, but did positively correlate with rainfall (**Supporting Information Fig. S6**). Phosphate concentrations increased in the summer compared to the winter in the nearshore and transition clusters, while silicate concentrations increased in the summer compared to the winter in the nearshore cluster (**Supporting Information Table S9**).

Ratios of phytoplankton pigments to Tchla representing cyanobacteria, diatoms, and pelagophytes showed correlations with seawater temperature in the nearshore community type, suggestive of seasonality in the abundance of major phytoplankton groups (**Supporting Information Fig. S7**). Correlations between seawater temperature and pigment to Tchla ratios were not detected in the transition and offshore cluster (**Supporting Information Fig. S7**).

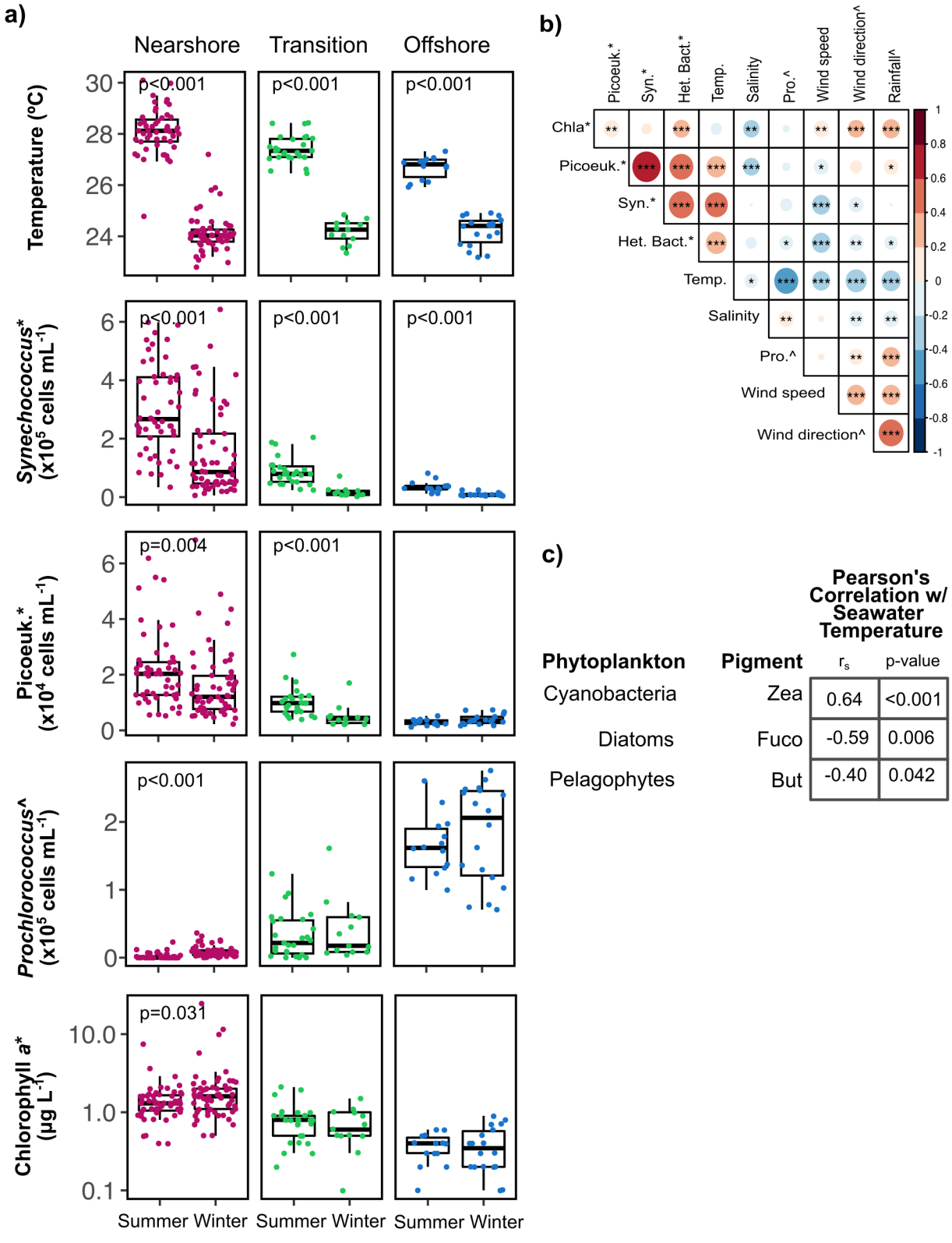
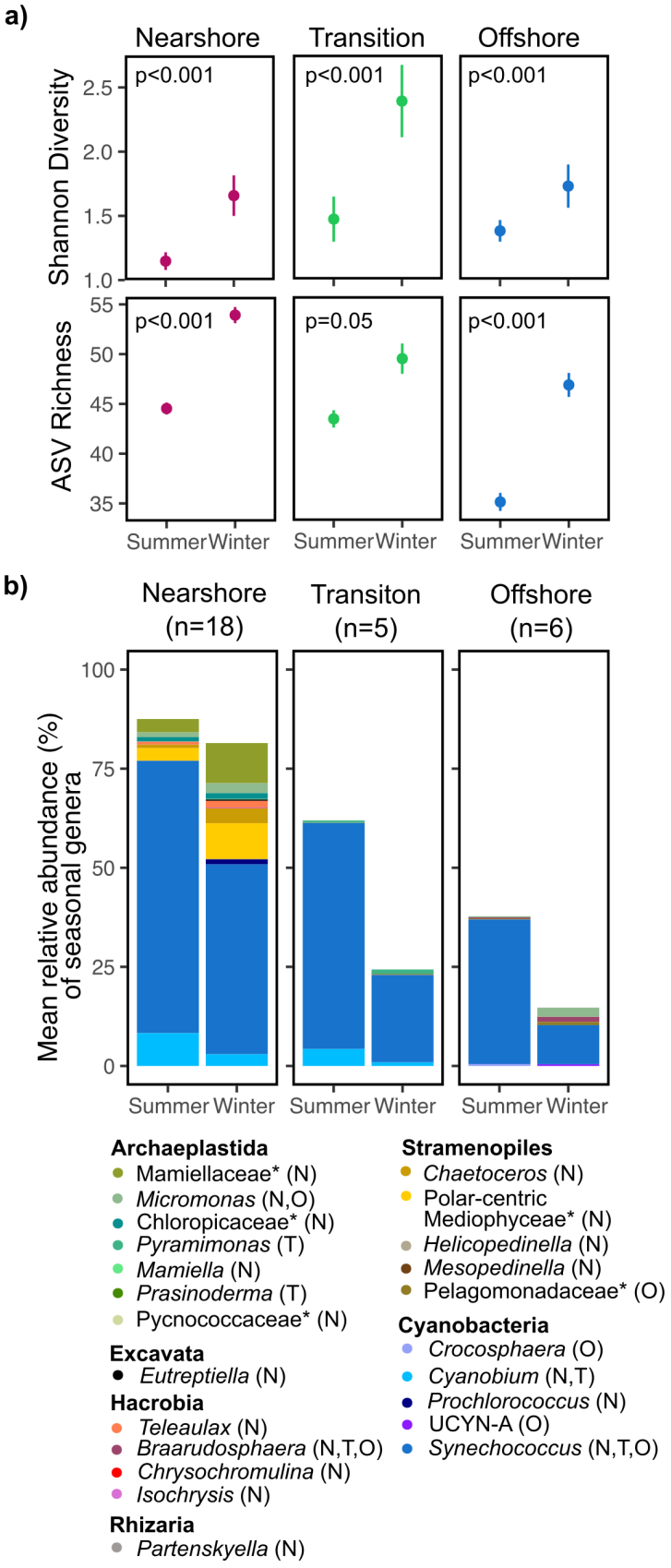


Fig. 6. a) Changes in seawater temperature, cell counts of *Synechococcus*, photosynthetic

469 picoeukaryotes, *Prochlorococcus*, and chlorophyll *a* across seasons and environments. b)
470 Correlations between chlorophyll *a* concentrations and other environmental parameters in the
471 nearshore. c) Significant correlations between phytoplankton pigment:tchla ratios and seawater
472 temperature in the nearshore. An asterisk (*) denotes variables with log transformations, while a
473 carrot (^) denotes variables with log+1 transformations. Het.Bac: heterotrophic bacteria (cells
474 mL^{-1}); Syn: *Synechococcus* (cells mL^{-1}); Pro: *Prochlorococcus* (cells mL^{-1}); Picoeuk:
475 Photosynthetic picoeukaryotes (cells mL^{-1}); Temp: Seawater temperature (°C); Chla: Chlorophyll
476 *a* ($\mu g\ L^{-1}$); Salinity (ppt); Wind direction (degrees); Wind Speed (ms^{-1}); Rainfall (mm); Zea:
477 Zeaxanthin; Fuco: Fucoxanthin; But: 19′-butanoyloxyfucoxanthin.
478
479 Seasonal differences were found in phytoplankton alpha diversity and the relative
480 abundance of phytoplankton genera across the nearshore, transition, and offshore. ASV richness
481 and Shannon diversity increased in the winter compared to the summer in all three community
482 types (**Fig. 7a**, **Supporting Information Table S10**). Seasonality was observed among 23
483 phytoplankton genera (or groups unclassified at the genus-level, but classified at the family-
484 level) within each of the clusters, including 18 nearshore, 5 in the transition, and 6 offshore
485 (**Fig.7b, Supporting Information Table S11**). Seasonal genera accounted for 84.9±7.4,
486 48.6±18.9, and 23.2±13.1% of the relative abundance of the community on average in the
487 nearshore, transition, and offshore clusters, respectively. *Synechococcus* was the most abundant
488 seasonal genus, increasing in relative abundance in the summer months across all community
489 type clusters (**Fig.7b,c**). *Cyanobium*, *Crocosphaera*, and unidentified Pycnococcaceae also
490 increased in abundance in the summer months (**Fig. 7b**). Seasonal genera more often increased in
491 relative abundance during the winter (n=18) including those belonging to diatoms (e.g.

492 *Chaetoceros* and unclassified polar-centric Mediophyceae), green algae (e.g. unclassified
493 Mamiellophyceae, *Micromonas*, and *Mamiella*), prymnesiophytes (e.g. *Isochrysis*,
494 *Chrysochromulina*, and *Braarudosphaera*), and Dictyochophyceae (*Helicopedinella*,
495 *Mesopedinella*) (**Fig. 7b**). In the nearshore, the dominant phytoplankton in the winter remained
496 Cyanobacteria followed by Mamiellophyceae and then diatoms, although at times both
497 Mamiellophyceae and diatoms exceeded >30% of the total phytoplankton relative abundance
498 (**Fig. 7c**).



499

Fig. 7. a) ASV richness and Shannon's diversity across all three community types differ during the summer and winter seasons. b) Mean relative abundances of the 29 significantly seasonal genera across the three community types for summer and winter seasons. c) Nearshore seasonal phytoplankton genera from Cyanobacteria, Stramenopiles, and Archaeplastida with relative abundances >0.5% on average. A local polynomial regression fit line is shown for each genus. Phytoplankton groups that were classified at the family-level but unidentified at the genus-level are denoted with an asterisk.

Discussion

The activity and distribution of phytoplankton has important implications for food web dynamics and biocultural restoration and management of near-island waters of the tropical Pacific, especially under predicted climate change conditions that are expected to limit nutrient availability, phytoplankton size, productivity, and biomass in open oceans. Ecosystem-level time-series analyses spanning estuarine waters within Kāne'ohe Bay of the Hawaiian island of O'ahu to the North Pacific Subtropical Gyre revealed that surface ocean biogeochemistry, phytoplankton biomass, and phytoplankton community structure varied dramatically across both broad (e.g. nearshore Kāne'ohe Bay to the NPSG) and narrow (e.g. nearshore Kāne'ohe Bay to adjacent offshore waters) spatial scales. Through investigations of these spatiotemporal dynamics, we gained insight into the ecology and phenology of phytoplankton communities of surface oceans in the tropical Pacific and identified indicators to help evaluate deviations from current conditions.

523

524 *Spatially distinct phytoplankton and biogeochemical regimes within and adjacent to Kāne'ohe*
525 *Bay*

526 Our sampling approach provides an increased resolution upon satellite-based studies
527 (Gove et al. 2016; Messié et al. 2022), which are typically not reliable within ~5 km to shore due
528 to heavy cloud cover, along with a synoptic view of the biogeochemistry and phytoplankton
529 communities. Compared to offshore and open ocean waters, nearshore Kāne'ohe Bay had
530 elevated chlorophyll *a* concentrations- a 3-fold increase in phytoplankton biomass within 6 km
531 from coastal Kān'eohe Bay and an 18-fold increase over a roughly 100 km distance. Inorganic
532 nutrient and chlorophyll *a* concentrations increased with decreasing salinity, suggesting that
533 freshwater input from streams is an important driver of nutrient delivery and subsequent
534 phytoplankton enhancement within He'eia Fishpond and Kāne'ohe Bay. Previous studies within
535 Kāne'ohe Bay have also revealed stream-delivered nutrient input as an important driver of
536 phytoplankton enhancement (Yeo et al. 2013), as well as additional processes including
537 submarine groundwater discharge (McKenzie et al. 2019) and human-driven pollution (Ringuet
538 and Mackenzie 2005). Transport from offshore, subsurface nutrient-rich waters to nearshore
539 waters via internal waves is likely an important process contributing to phytoplankton
540 enhancement of surface waters in the larger Hawaiian archipelago (Gove et al., 2016), although
541 its prevalence and impact specifically within Kāne'ohe Bay remains unknown.

542 Across Kāne'ohe Bay and the adjacent open ocean, phytoplankton communities resolved
543 into three distinct groups that coincide with spatial differences in biogeochemistry. Pigments
544 indicative of diatoms and cyanobacteria revealed that these groups were the main phytoplankton
545 contributing to the enhanced chlorophyll *a* concentrations in the nearshore environment of

Kāne‘ohe Bay. Some phytoplankton that were more abundant closer to shore, such as *Teleaulax* (Cryptophyta), *Pyamimonas* (Chlorophyta), *Tetraselmis* (Chlorophyta), Chlorarachniophytes (Rhizaria), and dinoflagellates, are associated with mixotrophic lifestyles (Stoecker et al. 2016). Mixotrophy provides a crucial trophic link in planktonic food webs by supplementing primary production via heterotrophy, increasing carbon transfer to high trophic levels, and serving as a source of nutrients (Stoecker et al. 2016). While establishing the relative importance of mixotrophy across the nearshore to the adjacent offshore environment requires further investigation, these initial insights show distinctions in food web dynamics, phytoplankton ecologies, and contributors to primary productivity across the nearshore to open ocean waters of the tropical Pacific.

Drivers of seasonality in nearshore chlorophyll a *concentrations*

While seasonality within transition and offshore phytoplankton communities was muted, nearshore phytoplankton biomass was significantly elevated in winter, and members with the highest relative abundance in the nearshore phytoplankton community showed high seasonality. Chlorophyll *a* concentrations varied seasonally in nearshore Kāne‘ohe Bay, where it increased with wind speed, rainfall, and wind direction. In Hawai‘i, storm events generally increase during the winter months where storm-associated rainfall and wind may serve to elevate inorganic nutrient concentrations through increased stream outflow, leading to short-term enhancement of phytoplankton biomass in nearshore Kāne‘ohe Bay (Ringuet and Mackenzie 2005; Cox et al. 2006; Yeo et al. 2013). Periods of intense rainfall have strong impacts on the food web dynamics of Kāne‘ohe Bay, with wet periods decreasing trophic complexity and increasing total community biomass and the transfer of production to metazoans (Selph et al. 2018). Because

569 phytoplankton growth and loss processes can occur over short time scales of a day to a week, the
570 near-monthly sampling interval employed in KByT is unlikely to fully capture the rapid
571 fluctuations of chlorophyll *a* concentrations that occur in response to storm events. Thus, our
572 observations likely underestimate the episodic variability of phytoplankton biomass in the
573 nearshore (Yeo et al. 2013).

574 Distinct growth and nutrient uptake strategies between *Synechococcus* and diatoms, along
575 with changes in the environment that might favor one over the other, likely produced the
576 dynamic seasonal patterns observed in nearshore Kāne‘ohe Bay: summer months showed a near
577 doubling of the cellular and relative abundance of *Synechococcus*, while winter months were
578 marked by sharp increases in the relative abundances of diatoms such as *Chaetoceros* and
579 unidentified polar-centric Mediophyceae. To understand competition outcomes under different
580 environmental conditions, resource competition theory characterizes trade-offs between slow-
581 growing nutrient specialists with high affinity for uptake but low maximum growth (e.g.
582 *Prochlorococcus* and *Synechococcus*) and fast-growing nutrient opportunists with high
583 maximum growth but low affinity for uptake (e.g. flagellates and diatoms) (Dutkiewicz et al.
584 2009). Due to their small size and high uptake capacity, *Synechococcus* are likely less limited by
585 the typically low nitrogen concentrations within Kāne‘ohe Bay than other phytoplankton (Burson
586 et al. 2018). In addition, seawater temperatures have been shown to positively correlate with
587 *Synechococcus* cellular abundances, likely because seasonal increases in temperature positively
588 impact *Synechococcus* division rates (Hunter-Cevera et al. 2016). Thus, in the summer under
589 high-light, warm-water temperatures and limited nutrients, *Synechococcus* dominated Kāne‘ohe
590 Bay. In contrast, despite higher nutrient requirements due to their larger cell sizes, diatoms have
591 high rates of growth allowing them to outcompete other phytoplankton under periods of elevated

592 nutrient availability (Laws 1975). Pulses of nutrients from the more frequent storm events during
593 the winter could have led to the observed fluxes in diatom relative abundances.
594
595 *Phytoplankton indicators of climate change impacts*
596 Phytoplankton biodiversity (e.g. ASV richness, Shannon's diversity) in Kāne'ohe Bay
597 was significantly elevated in comparison to the adjacent offshore. Near-island environments can
598 export diversity offshore and, importantly, biodiversity can offer functional redundancy and
599 ecological stability in times of environmental perturbation (Messié et al. 2022). In a recent study
600 that modeled changes to phytoplankton biodiversity under climate change conditions, some
601 tropical regions were found to face up to 30% of phytoplankton types becoming locally
602 extirpated (Henson et al. 2021). Importantly, phytoplankton of higher size classes, predominantly
603 diatoms, are expected to be lost due to increased nutrient limitation (Flombaum et al. 2020).
604 Diatoms serve as an important part of the diet of herbivorous fish grown in Hawaiian aquaculture
605 systems (Hiatt 1947), and contribute significantly to primary productivity in nearshore systems
606 broadly. Diatoms also contribute a large portion of the total chlorophyll *a* in the nearshore
607 environment of Kāne'ohe Bay. Shifts or reduction in diatom abundance within the He'eia
608 Fishpond and Kāne'ohe Bay when compared the baseline knowledge characterized here might
609 thus help to identify impacts of ocean warming and stratification in advance of shifts in the food
610 web structure.
611 *Synechococcus* and *Prochlorococcus*, the small cyanobacteria that dominated the total
612 phytoplankton relative abundance in our study, may also provide valuable bioindicators of
613 environmental change within Kāne'ohe Bay under climate change scenarios. *Prochlorococcus*
614 and *Synechococcus* are responsible for roughly 25% of the ocean's net primary production

(Flombaum et al. 2013) and will likely further increase in abundance with projected climate change conditions (Flombaum et al. 2020). These marine cyanobacteria encompass fine-scale genetic diversity that distinguishes their ecologies, metabolisms, and biogeochemical roles at the level of major and minor clades (Berube et al. 2019). These clade identities are often difficult to resolve with the use of single gene markers, and thus metagenomic read recruitment can increase genetic resolution.

Across the open ocean NPSG to nearshore Kāneʻohe Bay, *Prochlorococcus* Clade HLII and *Synechococcus* Clade II (SC 5.1) were the most abundant, consistent with previous reports from oligotrophic oceans (Delmont and Eren 2018; Lee et al. 2019). *Prochlorococcus* Clade HLII from the offshore waters adjacent to Kāneʻohe Bay and Station ALOHA showed high similarity in population structure and thus may represent continuous populations with ongoing gene flow responding to similar environmental parameters in both environments. Continued examination of population structure, as well as cellular and relative abundances of these cyanobacteria, could identify eventual expansions of ultraoligotrophic waters into Kāneʻohe Bay, shifts in gene-flow, and selection for clades with unique ecological adaptations. Given the high seasonality of *Synechococcus* cellular and relative abundance in the nearshore environment, changes in the magnitude and timing of these metrics could also be used to identify alterations in seasonality and associated food-web dynamics within Kāneʻohe Bay.

Implications for biocultural restoration

The enhancement of phytoplankton biomass in the estuarine Heʻeia Fishpond and coastal Kāneʻohe Bay provides critical ecosystem services and is an important consideration for biocultural restoration activities, community-based research efforts, and resource management.

638 Within He‘eia and across the islands of Hawai‘i, biocultural practitioners are undertaking
639 restoration projects to maximize primary and secondary production of the estuarine environment,
640 including removing invasive mangrove, managing stream use to ensure adequate flow and water
641 quality, and engineering water exchange through repairs and updates to fishpond walls
642 (Möhlenkamp et al. 2019). The baseline understanding of phytoplankton communities provided
643 by this study helps to inform biocultural stewards of He‘eia and Kāne‘ohe Bay of the conditions
644 that promote phytoplankton growth and subsequently herbivorous fish growth.

645 High-resolution sampling was key to identifying the magnitude of chlorophyll *a*
646 enhancement across the system, areas of high localized phytoplankton biomass enhancement
647 such as He‘eia Fishpond, and drivers of chlorophyll *a* enhancement including freshwater input
648 and increased storm conditions and nutrient concentrations. Despite fine-scale biogeochemical
649 differences within Kāne‘ohe Bay, phytoplankton communities sampled from the He‘eia
650 Fishpond and across the northern, central, and southern sections of the bay grouped as one
651 nearshore community type. The similar phytoplankton communities found within the estuarine
652 fishpond and the nearshore environment of the bay emphasizes the connectedness of these
653 estuarine and coastal systems and highlights the need to manage them in close coordination.

654 The particularly wet conditions during the winter months appears to play a substantial
655 role in determining the variability of phytoplankton biomass and restructuring of the coastal food
656 web within the coastal environment studied here. Growing this baseline understanding of
657 phytoplankton cycling to relate to the timing and conditions documented in Hawaiian knowledge
658 systems like *kaulana mahina*, the Hawaiian lunar calendar (Nu‘uhiwa 2019), and in regard to life
659 cycles of bioculturally relevant species like ‘ama‘ama (*Mugil cephalus*) that feeds on

660 microphytoplankton at juvenile stages (Hiatt 1947), would further advance this area of study and
661 utility for management within Hawaiian aquaculture systems.

662 Despite a poor understanding of the outcomes, near-island food webs will likely shift
663 with ongoing climate change impacts on phytoplankton communities and biomass. Collaborating
664 across diverse knowledge systems could improve the ability for local people to document and
665 adapt to changes in near-island marine resources (Winter et al. 2020b). Our collaboratively
666 developed study reveals distinct spatial and seasonal dynamics that define phytoplankton
667 communities and biogeochemical conditions from an estuarine aquaculture system on the coast
668 of O‘ahu, Hawai‘i, to the open ocean of the North Pacific Subtropical Gyre. Understanding the
669 seasonal and spatial dynamics underlying phytoplankton communities and biogeochemistry of
670 near-island environments in the tropical Pacific provides the necessary knowledge to further co-
671 develop capacities to model and track changes to the marine food web, and to build resilience
672 now and in the future.

673 **Acknowledgements**

674 We thank Paepae o He‘eia and the He‘eia National Estuarine Research Reserve for their
675 involvement in and perspectives on the directions on this research. We thank Jason Jones, Ciara
676 Ratum, Hanako Mochimaru, Oscar Ramfelt, Amber Boettiger, Clarisse E.S. Sullivan, Elizabeth
677 A. Monaghan, and Evan Freel for their assistance collecting samples, Gus Robertson for
678 assistance with sonde calibration, Andrew McGowan for his help with the creation of maps, and
679 Karen Selph for flow cytometry measurements. Any opinions, findings, and conclusions or
680 recommendations expressed in this material are those of the author(s) and do not necessarily
681 reflect the views of the National Science Foundation or the National Oceanic and Atmospheric
682 Administration (NOAA). This research was funded by the National Oceanic and Atmospheric
683 Administration (NOAA) Margaret A. Davidson Fellowship (No. NA20NOS4200123), National
684 Science Foundation Graduate Research Fellowship Program under Grant No. 1842402, the
685 Colonel Willys E. Lord, DVM and Sadina L. Lord Scholarship, and the University of Hawai‘i at
686 Mānoa Ecology, Evolution, and Conservation Biology’s Maybelle Roth ARCS Award to SJ
687 Tucker and NSF grants OCE-1538628 and OCE- 2149128 to MS Rappé. This is SOEST
688 contribution XXX and HIMB contribution XXX.

689 **Citations and References**

690

691 Aitchison, J. 1982. The Statistical Analysis of Compositional Data. J Royal Statistical Soc Ser B
692 Methodol **44**: 139–160. doi:10.1111/j.2517-6161.1982.tb01195.x
693 Auladell, A., A. Barberán, R. Logares, E. Garcés, J. M. Gasol, and I. Ferrera. 2021. Seasonal
694 niche differentiation among closely related marine bacteria. Isme J 1–12. doi:10.1038/s41396-
695 021-01053-2
696 Azam, F., T. Fenchel, J. Field, J. Gray, L. Meyer-Reil, and F. Thingstad. 1983. The Ecological
697 Role of Water-Column Microbes in the Sea. Mar Ecol Prog Ser **10**: 257–263.
698 doi:10.3354/meps010257
699 Berube, P. M., A. Rasmussen, R. Braakman, R. Stepanauskas, and S. W. Chisholm. 2019.
700 Emergence of trait variability through the lens of nitrogen assimilation in Prochlorococcus.
701 Elife **8**. doi:10.7554/elife.41043
702 Bidigare, R., O. Schofield, and B. Prezelin. 1989. Influence of zeaxanthin on quantum yield of
703 photosynthesis of Synechococcus clone WH7803 (DC2). Mar Ecol Prog Ser **56**: 177–188.
704 doi:10.3354/meps056177
705 Bolyen, E., J. R. Rideout, M. R. Dillon, and others. 2019. Reproducible, interactive, scalable and
706 extensible microbiome data science using QIIME 2. Nat Biotechnol **37**: 852–857.
707 doi:10.1038/s41587-019-0209-9
708 Burson, A., M. Stomp, E. Greenwell, J. Grosse, and J. Huisman. 2018. Competition for nutrients
709 and light: testing advances in resource competition with a natural phytoplankton community.
710 Ecology **99**: 1108–1118. doi:10.1002/ecy.2187

711 Calbet, A., and M. R. Landry. 1999. Mesozooplankton influences on the microbial food web:
712 Direct and indirect trophic interactions in the oligotrophic open ocean. Limnol. Oceanogr. **44**:
713 1370–1380. doi:10.4319/lo.1999.44.6.1370
714 Callahan, B. J., J. Wong, C. Heiner, S. Oh, C. M. Theriot, A. S. Gulati, S. K. McGill, and M. K.
715 Dougherty. 2019. High-throughput amplicon sequencing of the full-length 16S rRNA gene
716 with single-nucleotide resolution. **4**: e2492-28. doi:10.1101/392332
717 Carswell, T., M. Costa, E. Young, N. Komick, J. Gower, and R. Sweeting. 2017. Evaluation of
718 MODIS-Aqua Atmospheric Correction and Chlorophyll Products of Western North American
719 Coastal Waters Based on 13 Years of Data. Remote Sensing **9**: 1063–24.
720 doi:10.3390/rs9101063
721 Cox, E. F., M. Ribes, and I. K. RA. 2006. Temporal and spatial scaling of planktonic responses
722 to nutrient inputs into a subtropical embayment. Mar. Ecol. Prog. Ser. **324**: 19–35.
723 doi:10.3354/meps324019
724 Delmont, T. O., and A. M. Eren. 2018. Linking pangenomes and metagenomes: the
725 Prochlorococcus metapangenome. PeerJ **6**: e4320. doi:10.7717/peerj.4320
726 Doré, H., G. K. Farrant, U. Guyet, and others. 2020. Evolutionary Mechanisms of Long-Term
727 Genome Diversification Associated With Niche Partitioning in Marine Picocyanobacteria.
728 Front Microbiol **11**: 567431. doi:10.3389/fmicb.2020.567431
729 Doty, M. S., and M. Oguri. 1956. The island mass effect. J. Cons. perm. int. Explor. Mer **22**: 33–
730 37. doi:10.1093/icesjms/22.1.33
731 Dutkiewicz, S., M. J. Follows, and J. G. Bragg. 2009. Modeling the coupling of ocean ecology
732 and biogeochemistry. Glob. Biogeochem. Cycles **23**: n/a-n/a. doi:10.1029/2008gb003405

733 Eddy, T. D., J. R. Bernhardt, J. L. Blanchard, and others. 2021. Energy Flow Through Marine
734 Ecosystems: Confronting Transfer Efficiency. Trends Ecol. Evol. **36**: 76–86.
735 doi:10.1016/j.tree.2020.09.006
736 Eren, A. M., E. Kiefl, A. Shaiber, and others. 2021. Community-led, integrated, reproducible
737 multi-omics with anvi'o. Nat Microbiol **6**: 3–6. doi:10.1038/s41564-020-00834-3
738 Eren, A. M., J. H. Vineis, H. G. Morrison, and M. L. Sogin. 2013. A Filtering Method to
739 Generate High Quality Short Reads Using Illumina Paired-End Technology. Plos One **8**:
740 e66643. doi:10.1371/journal.pone.0066643
741 Finley, A., B. Sudipto, and Hjelle. 2017. MBA: Multilevel B-Spline Approximation,.
742 Flombaum, P., J. L. Gallegos, R. A. Gordillo, and others. 2013. Present and future global
743 distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. Proc National
744 Acad Sci **110**: 9824–9829. doi:10.1073/pnas.1307701110
745 Flombaum, P., W.-L. Wang, F. W. Primeau, and A. C. Martiny. 2020. Global picophytoplankton
746 niche partitioning predicts overall positive response to ocean warming. Nat. Geosci. **13**: 116–
747 120. doi:10.1038/s41561-019-0524-2
748 Galili, T. 2015. dendextend: an R package for visualizing, adjusting and comparing trees of
749 hierarchical clustering. Bioinformatics **31**: 3718–3720. doi:10.1093/bioinformatics/btv428
750 Gove, J. M., M. A. McManus, A. B. Neuheimer, and others. 2016. Near-island biological
751 hotspots in barren ocean basins. Nature Communications **7**: 10581–8.
752 doi:10.1038/ncomms10581
753 Guillou, L., D. Bachar, S. Audic, and others. 2013. The Protist Ribosomal Reference database
754 (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated
755 taxonomy. Nucleic Acids Res **41**: D597–D604. doi:10.1093/nar/gks1160

Henson, S. A., B. B. Cael, S. R. Allen, and S. Dutkiewicz. 2021. Future phytoplankton diversity in a changing climate. Nat Commun **12**: 5372. doi:10.1038/s41467-021-25699-w

Hiatt, R. 1947. Food-Chains and the Food Cycle in Hawaiian Fish Ponds.–Part I. The Food and Feeding Habits of Mullet (Mugil Cephalus), Milkfish (Chanos Chanos), and the Ten-Pounder (Elops Machnata). Trans. Am. Fish. Soc. 250–261. doi:10.1577/1548-8659(1944)74[250:fatfci]2.0.co;2

Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous Inference in General Parametric Models. Biometrical Journal **50**: 346–363. doi:10.1002/bimj.200810425

Hunter-Cevera, K. R., M. G. Neubert, R. J. Olson, A. R. Solow, A. Shalapyonok, and H. M. Sosik. 2016. Physiological and ecological drivers of early spring blooms of a coastal phytoplankter. Science **354**: 326–329. doi:10.1126/science.aaf8536

Hyatt, D., G.-L. Chen, P. F. Locascio, M. L. Land, F. W. Larimer, and L. J. Hauser. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11**: 119. doi:10.1186/1471-2105-11-119

Jokiel. 1991. Jokiel's illustrated scientific guide to K¯ane'ohe Bay. 1–66. doi:10.13140/2.1.3051.9360

Karl, D. M., and M. J. Church. 2014. Microbial oceanography and the Hawaii Ocean Time-series programme. Nature Reviews Microbiology **12**: 699–713. doi:10.1038/nrmicro3333

Kelly, M. 1973. Some legendary and historical aspects of Heeia Fishpond, Koolau, Oahu, Bishop Museum.

Kolde, R. 2019. pheatmap: Pretty Heatmaps. R package version 1.0.12,.

777 Kwiatkowski, L., O. Aumont, and L. Bopp. 2018. Consistent trophic amplification of marine
778 biomass declines under climate change. Glob Change Biol **25**: 218–229.
779 doi:10.1111/gcb.14468
780 Lahti, and S. Shetty. 2017. Tools for microbiome analysis in R.
781 Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat Meth **9**:
782 357–359. doi:10.1038/nmeth.1923
783 Laws, E. A. 1975. The Importance of Respiration Losses in Controlling the Size Distribution of
784 Marine Phytoplankton. Ecology **56**: 419–426. doi:10.2307/1934972
785 Lee, M. D. 2019. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics **35**:
786 btz188. doi:10.1093/bioinformatics/btz188
787 Lee, M. D., N. A. Ahlgren, J. D. Kling, N. G. Walworth, G. Rocap, M. A. Saito, D. A. Hutchins,
788 and E. A. Webb. 2019. Marine Synechococcus isolates representing globally abundant
789 genomic lineages demonstrate a unique evolutionary path of genome reduction without a
790 decrease in GC content. Environmental Microbiology. doi:10.1111/1462-2920.14552
791 Li, G., L. Cheng, J. Zhu, K. E. Trenberth, M. E. Mann, and J. P. Abraham. 2020. Increasing
792 ocean stratification over the past half-century. Nat Clim Change **10**: 1116–1123.
793 doi:10.1038/s41558-020-00918-2
794 Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion
795 for RNA-seq data with DESeq2. Genome Biol. **15**: 31–21. doi:10.1186/s13059-014-0550-8
796 Lowe, R. J., J. L. Falter, S. G. Monismith, and M. J. Atkinson. 2009. A numerical study of
797 circulation in a coastal reef-lagoon system. Journal of Geophysical Research **114**: 997.
798 doi:10.1029/2008jc005081

799 Mantoura, R. F. C., and C. A. Llewellyn. 1983. The rapid determination of algal chlorophyll and
800 carotenoid pigments and their breakdown products in natural waters by reverse-phase high-
801 performance liquid chromatography. Anal Chim Acta **151**: 297–314. doi:10.1016/s0003-
802 2670(00)80092-6
803 McKenzie, T., H. Dulai, and J. Chang. 2019. Parallels between stream and coastal water quality
804 associated with groundwater discharge. PLoS ONE **14**: e0224513.
805 doi:10.1371/journal.pone.0224513
806 McMurdie, P. J., and S. Holmes. 2013. phyloseq: an R package for reproducible interactive
807 analysis and graphics of microbiome census data. M. Watson [ed.]. PLoS ONE **8**: e61217.
808 doi:10.1371/journal.pone.0061217
809 Mende, D. R., J. A. Bryant, F. O. Aylward, J. M. Eppley, T. Nielsen, D. M. Karl, and E. F.
810 DeLong. 2017. Environmental drivers of a microbial genomic transition zone in the ocean's
811 interior. Nat Microbiol **2**: 1367–1373. doi:10.1038/s41564-017-0008-3
812 Messié, M., A. Petrenko, A. M. Doglioli, E. Martinez, and S. Alvain. 2022. Basin-scale
813 biogeochemical and ecological impacts of islands in the tropical Pacific Ocean. Nat Geosci 1–
814 6. doi:10.1038/s41561-022-00957-8
815 Minoche, A. E., J. C. Dohm, and H. Himmelbauer. 2011. Evaluation of genomic high-throughput
816 sequencing data generated on Illumina HiSeq and Genome Analyzer systems. Genome Biol
817 **12**: R112. doi:10.1186/gb-2011-12-11-r112
818 Möhlenkamp, P., C. Beebe, M. McManus, A. Kawelo, K. Kotubetey, M. Lopez-Guzman, C.
819 Nelson, and R. Alegado. 2019. Kū Hou Kuapā: Cultural restoration improves water budget
820 and water quality dynamics in He'eia Fishpond. Sustainability **11**: 161–25.
821 doi:10.3390/su11010161

822 Monger, B. C., and M. R. Landry. 1993. Flow cytometric analysis of marine bacteria with
823 Hoechst 33342. Applied and Environmental Microbiology **59**: 905–911.
824 doi:10.1128/aem.59.3.905-911.1993
825 Moore, J. K., W. Fu, F. Primeau, and others. 2018. Sustained climate warming drives declining
826 marine biological productivity. Science **359**: 1139–1143. doi:10.1126/science.aao6379
827 Nu‘uhiwa, K. 2019. Papakū makawalu: A Methodology and Pedagogy of Understanding the
828 Hawaiian Universe, p. 39–49. *In* N. Wilson-Hokowhitu [ed.], The Past Before Us :
829 Mookauhau As Methodology. University of Hawaii Press.
830 Parada, A. E., D. M. Needham, and J. A. Fuhrman. 2016. Every base matters: assessing small
831 subunit rRNA primers for marine microbiomes with mock communities, time series and
832 global field samples. Environmental Microbiology **18**: 1403–1414. doi:10.1111/1462-
833 2920.13023
834 Pritchard, L., R. H. Glover, S. Humphris, J. G. Elphinstone, and I. K. Toth. 2015. Genomics and
835 taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal.
836 Methods **8**: 12–24. doi:10.1039/c5ay02550h
837 Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner.
838 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-
839 based tools. Nucleic Acids Res **41**: D590–D596. doi:10.1093/nar/gks1219
840 Redfield, A. C. 1960. The biological control of chemical factors in the environment. Sci Prog **11**:
841 150–70. doi:10.1086/646891
842 Ringuet, S., and F. T. Mackenzie. 2005. Controls on nutrient and phytoplankton dynamics during
843 normal flow and storm runoff conditions, southern Kaneohe Bay, Hawaii. Estuaries **28**: 327–
844 337. doi:10.1007/bf02693916

845 Ruf, T. 1999. The Lomb-Scargle Periodogram in Biological Rhythm Research: Analysis of
846 Incomplete and Unequally Spaced Time-Series. Biol Rhythm Res **30**: 178–201.
847 doi:10.1076/brhm.30.2.178.1422
848 Selph, K., E. Goetze, M. Jungbluth, P. Lenz, and G. Kolker. 2018. Microbial food web
849 connections and rates in a subtropical embayment. Mar Ecol Prog Ser **590**: 19–34.
850 doi:10.3354/meps12432
851 Smith, S. V. 1981. Responses of Kaneohe Bay, Hawaii, to Relaxation of Sewage Stress, p. 391–
852 410. *In* Estuaries and Nutrients. Estuaries and Nutrients. doi:10.1007/978-1-4612-5826-1_18
853 Stock, C. A., J. G. John, R. R. Rykaczewski, and others. 2017. Reconciling fisheries catch and
854 ocean productivity. Proc National Acad Sci **114**: E1441–E1449.
855 doi:10.1073/pnas.1610238114
856 Stoecker, D. K., P. J. Hansen, D. A. Caron, and A. Mitra. 2016. Mixotrophy in the Marine
857 Plankton. Annu Rev Mar Sci **9**: 1–25. doi:10.1146/annurev-marine-010816-060617
858 Strimmer, K. 2008. fdrtool: a versatile R package for estimating local and tail area-based false
859 discovery rates. Bioinformatics **24**: 1461–1462. doi:10.1093/bioinformatics/btn209
860 Suzuki, M. T., O. Béjà, L. T. Taylor, and E. F. DeLong. 2001. Phylogenetic analysis of
861 ribosomal RNA operons from uncultivated coastal marine bacterioplankton. Environmental
862 Microbiology **3**: 323–331. doi:10.1046/j.1462-2920.2001.00198.x
863 Tucker, S. J., K. C. Freel, E. A. Monaghan, C. E. S. Sullivan, O. Ramfelt, Y. M. Rii, and M. S.
864 Rappé. 2021. Spatial and temporal dynamics of SAR11 marine bacteria across a nearshore to
865 offshore transect in the tropical Pacific Ocean. Peerj **9**: e12274. doi:10.7717/peerj.12274

866 Utter, D. R., G. G. Borisy, A. M. Eren, C. M. Cavanaugh, and J. L. M. Welch. 2020.
867 Metapangenomics of the oral microbiome provides insights into habitat adaptation and
868 cultivar diversity. **1**: 808410–40. doi:10.1101/2020.05.01.072496
869 Wei, T., and V. Simko. 2021. corrplot: A visualization of a correlation matrix.
870 Welschmeyer, N. A. 1994. Fluorometric analysis of chlorophyll a in the presence of chlorophyll
871 b and pheopigments. Limnol. Oceanogr. **39**: 1985–1992. doi:10.4319/lo.1994.39.8.1985
872 Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag New York,.
873 Willis, A. D., and B. D. Martin. 2020. Estimating diversity in networked ecological
874 communities. Biostatistics **23**: kxaa015. doi:10.1093/biostatistics/kxaa015
875 Winter, K., N. Lincoln, F. Berkes, and others. 2020a. Ecomimicry in Indigenous resource
876 management: optimizing ecosystem services to achieve resource abundance, with examples
877 from Hawai‘i. Ecol. Soc. **25**. doi:10.5751/es-11539-250226
878 Winter, K., Y. Rii, F. Reppun, and others. 2020b. Collaborative research to inform adaptive
879 comanagement: a framework for the He‘eia National Estuarine Research Reserve. Ecol Soc
880 **25**. doi:10.5751/es-11895-250415
881 Wright, E. 2016. Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. The R
882 Journal.
883 Xenopoulos, M. A., J. A. Downing, M. D. Kumar, S. Menden-Deuer, and M. Voss. 2017.
884 Headwaters to oceans: Ecological and biogeochemical contrasts across the aquatic continuum
885 M. Xenopoulos, J.A. Downing, M.D. Kumar, S. Menden-Deuer, and M. Voss [eds.]. Limnol.
886 Oceanogr. **62**: S3–S14. doi:10.1002/lno.10721

887 Yeo, S. K., M. J. Huggett, A. Eiler, and M. S. Rappé. 2013. Coastal bacterioplankton community
888 dynamics in response to a natural disturbance D.L. Kirchman [ed.]. PLoS ONE **8**: e56207-14.
889 doi:10.1371/journal.pone.0056207

890

891 **Data availability statement**

892 Data supporting the results within the manuscript are available within the main text (See
893 Materials & Methods). Sequencing data are available in the National Center for Biotechnology
894 Information (NCBI) Sequence Read Archive (SRA) under BioProject number PRJNA706753 as
895 well as PRJNA971314. Environmental data were submitted to BCO-DMO under
896 https://www.bco-dmo.org/project/663665. Code used in the analysis is available is available at
897 https://github.com/tucker4/Tucker_Phytoplankton_KByT_HeNERR.

898

899 **Author contribution statement**

900 SJT led the formal analyses and wrote the initial manuscript draft. SJT, YMR, KFC, MSR
901 collected the samples with assistance from AHK and KK. SJT, KFC, YMR, and MSR processed
902 samples. All authors were involved in the conceptualization, development of methodology,
903 interpretation of results, and providing input to the manuscript draft and revisions.