Environmental Microbiology

ENVIRONMENTAL MICROBIOLOGY

RESEARCH ARTICLE

Drivers of Marine Phytoplankton Diversity and Connectivity in the Galápagos Archipelago Spanning an ENSO Cycle

Prisca Lim[1] | Harvey Seim[1] | Oliva Torano[1] | Erika Neave[1,2] | Se Hyeon Jang[1,3] | Zackary Johnson[4] | Sara Haines[1] | Scott Gifford[1] | Natalie Cohen[1,5] | Carly M. Moreno[1,6] | Margarita Lankford[1] | Cristina Vintimilla Palacios[7] | Adrian Marchetti[1]

[1]Department of Earth, Marine and Environmental Sciences, University of North Carolina at Chapel Hill, Chapel Hill, USA | [2]Marine Management Organisation, Newcastle upon Tyne, UK | [3]Department of Oceanography, Chonnam National University, Gwangju, South Korea | [4]Nicholas School of the Environment. Biology, Civil & Environmental Engineering and Duke Microbiome Center, Duke University, North Carolina, USA | [5]Skidaway Institute of Oceanography, University of Georgia, Georgia, USA | [6]Marine Microbiomics Lab, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates | [7]Galápagos Science Center, Universidad San Francisco de Quito (USFQ), University of North Carolina at Chapel Hill (UNC), Chapel Hill, USA

Correspondence: Adrian Marchetti (amarchetti@unc.edu)

Received: 10 October 2024 | **Revised:** 20 June 2025 | **Accepted:** 28 June 2025

Funding: This work was supported by College of Arts and Sciences, University of North Carolina National Science Foundation (DEB2326027), (OCE1751805) UNC Center for Galapágos Studies.

Keywords: El Niño-southern oscillation (ENSO) | equatorial undercurrent (EUC) | Galapagos archipelago | phytoplankton ecology | upwelling

ABSTRACT

The Galápagos Islands are a biodiversity hotspot, largely due to the Equatorial Undercurrent (EUC) which supplies nutrient-rich waters to the euphotic zone and supports enhanced levels of primary productivity performed by phytoplankton. Understanding phytoplankton responses to changing environmental conditions is crucial for regional conservation and management efforts. Research cruises conducted between 2014 and 2022, spanning a major El Niño event in 2015 and a La Niña event in 2022, observed varying oceanic conditions and diverse phytoplankton community composition. At most EUC-influenced stations, larger-sized phytoplankton groups ($\geq 5\,\mu m$) were dominant while warmer, oligotrophic sites favoured smaller-sized phytoplankton groups ($< 5\,\mu m$). Predictably, nutrient supply was suppressed during the El Niño event associated with the weakening of the EUC and deepening of the thermocline. Counterintuitively, nutrient levels were not significantly enhanced during the La Niña event likely because increased stratification between the mixed and deep water layers reduced entrainment, particularly at Eastern stations. Protist community composition was evaluated using 18S rRNA gene metabarcoding; the majority of detected OTUs were associated with upwelling conditions prevalent around the archipelago. Taxonomic variability reflected heterogeneous environmental conditions generated by the convergence of multiple ocean currents. These results highlight the dynamic interplay of physical and biological factors shaping primary productivity in the Galápagos marine ecosystem.

1 | Introduction

The Galápagos Islands lie along the equator, roughly 1000 km west of the Ecuadorian coast. The region is wellknown for its unique biodiversity leading to the development of the Galápagos Marine Reserve (GMR) in 1998 to protect marine life. The high levels of endemic species in the GMR are the result of the convergence of multiple ocean currents at the archipelago, bringing together various water masses around the islands. Three major ocean currents (Figure 1) play distinct

roles in shaping the environmental conditions of the GMR, bringing waters of different temperatures, densities and varying levels of nutrients together (Fiedler and Talley 2006; Lindley and Barber 1998). The South Equatorial Current (SEC) is a westward-flowing surface current, fed by the Peruvian coastal upwelling and equatorial upwelling (Pennington et al. 2006). The North Equatorial Countercurrent (NECC) flows eastward at the surface, transporting waters from the western Pacific warm pool north of the SEC. Highest nutrient concentrations are supplied by the eastward-flowing Equatorial Undercurrent (EUC) that meets the western side of the Galápagos archipelago. It flows below the surface layer (Kessler 2006), around and through the archipelago, carrying nutrient-rich subtropical underwater which supports high

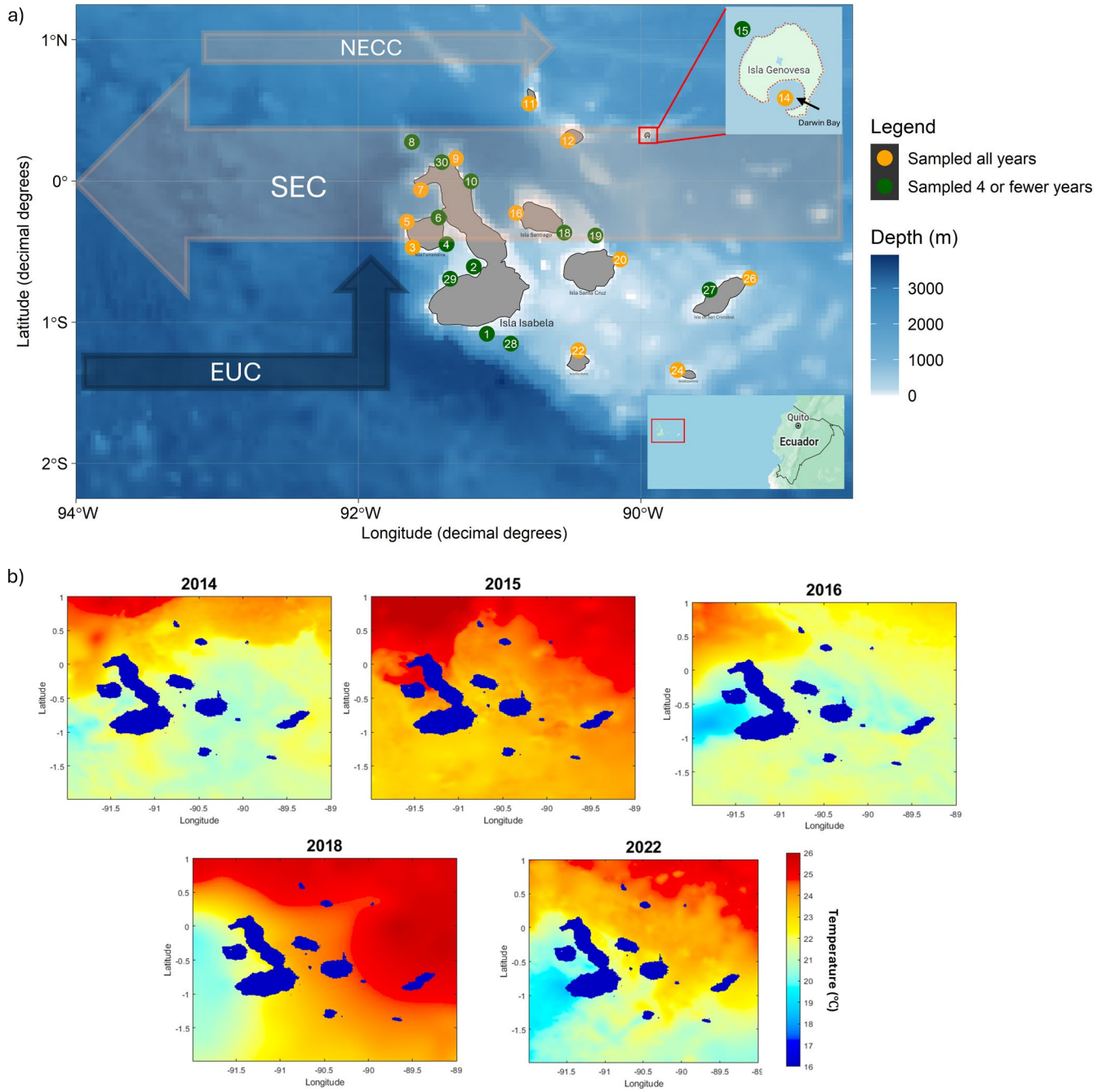

FIGURE 1 | (a) Bathymetry map of the Galápagos Archipelago indicating sampled stations. Yellow stations were sampled all five years whereas green stations were sampled four or fewer years. Approximate direction and position of main ocean currents are overlayed with blue arrow representing subsurface Equatorial Undercurrent (EUC), and orange arrows representing surface currents—South Equatorial Current (SEC) and North Equatorial Countercurrent (NECC). Inset of Isla Genovesa displays topography around station 14, located in the centre of a submerged crater known as Darwin Bay. All stations west of Isla Isabela are classified as Equatorial Undercurrent (EUC) stations due to the direct impact of upwelling waters. The remaining stations east of Isla Isabela are classified as the Eastern Region (ER). (b) Sea surface temperature plots showing average surface water temperatures around the archipelago during observation periods for each year. Data obtained from GHRSST Level 4 MUR Global Foundation Sea Surface Temperature Analysis (v4.1) (US NASA Jet Propulsion Laboratory Physical Oceanography Distributed Active Archive Center (JPL PO.DAAC) 2011).

primary productivity in the region (Chavez and Brusca 1991; Sakamoto et al. 1998).

The El Niño Southern Oscillation (ENSO) is a climate cycle with major impacts on climate patterns, particularly in the central and eastern equatorial Pacific Ocean, where the Galápagos Islands are located. During ENSO-neutral conditions, the EUC brings cold, nutrient-rich waters from the deep ocean to the surface surrounding the archipelago. These nutrients fuel phytoplankton growth and relatively high primary productivity compared to surrounding Equatorial Pacific waters, allowing the Galápagos Islands to develop into the unique biodiversity hotspot it is today. Under La Niña conditions, the upwelling of the EUC is intensified and thus higher levels of nutrients are expected to reach surface waters. However, under El Niño conditions, reversal of trade winds results in the weakening of the EUC and a coinciding deepening of the thermocline (Marin Jarrin and Lippmann 2019), and consequently lower amounts of nutrients being transported to the sunlit regions of the oceans to support primary productivity.

Phytoplankton contribute the majority of the net primary production in the Galápagos Islands, particularly since the steep topography around the islands limits the penetration of light to the seafloor required for the growth of macroalgae and seagrasses (Pessarrodona et al. 2022). Thus, marine phytoplankton directly or indirectly affect all organisms living in and around the archipelago. These microorganisms are highly sensitive to changes in oceanographic conditions, with their composition, biomass and productivity shifting with ENSO cycles and other seasonal variations. Environmental conditions around the Galápagos vary spatiotemporally due to the convergence of multiple ocean currents in the region, creating heterogeneous ecological niches that support extremely diverse plankton communities (Jang et al. 2022; Neave et al. 2021). However, little is known about how phytoplankton respond to these changes on both spatial and temporal scales. Interactions between microbes, including phytoplankton, can also play a significant role in structuring ocean ecosystems and carbon cycling (Worden et al. 2015). These interactions include symbiosis, parasitism, predation and mixotrophy, all of which affect how organic matter is being cycled throughout the water column.

While most of the oceans are warming due to climate change, parts of the Galápagos Islands have been observed to experience a cooling trend over the last several decades, due to the strengthening of the EUC (Karnauskas 2022). Consequently, the archipelago may serve as a climate refuge to organisms migrating away from warming waters. Given that phytoplankton compose the base of the marine food web and support all higher trophic level taxa, understanding how they respond under a variety of conditions, including ENSO cycles, will allow better prediction of how marine ecosystems will respond to a changing climate and can help guide conservation and management efforts within the GMR.

Few studies have characterised the phytoplankton composition and distribution around the Galápagos Islands. Jimenez (1981) utilised light microscopy to examine phytoplankton assemblages around the Galápagos archipelago and found that the diatom genera *Thalassiosira*, *Nitzchia* and *Chaetoceros* were important components of the phytoplankton community in upwelling regions, while flagellates predominated north and south of the equator. It is important to note that observations utilising light microscopy tend to be biased towards larger species that are easily discernible, such as diatoms and dinoflagellates. Accessory pigments have also been examined to assess phytoplankton composition in the region (McCulloch 2011), finding that relative abundances of diatoms and chlorophytes decreased during the 2004/05 El Niño event while cyanobacteria and haptophyte abundance increased. However, pigment-based analysis is unable to distinguish between different genera of phytoplankton within the same group. Compared to these techniques, 18S rRNA gene metabarcoding is advantageous in both sensitivity and specificity (De Vargas et al. 2015; Ebenezer et al. 2012). By targeting hypervariable regions within the 18S rRNA gene, the eukaryotic population can be resolved often to the genus level with high confidence.

This study aims to characterise the marine phytoplankton community around the Galápagos Islands and their responses to fluctuating environmental conditions. Measurements of physical, chemical and biological parameters in conjunction with 18S rRNA gene metabarcoding were employed to examine the phytoplankton community composition across three ENSO-neutral years (2014, 2016 and 2018), an El Niño event in 2015 and a La Niña event in 2022. Based on this dataset, three research questions were proposed: 1) How do key environmental variables—such as salinity, temperature and nutrient availability—vary across spatial zones and under different ENSO phases in the Galápagos Islands? 2) Do consistent protist community assemblages form under distinct environmental regimes, such as varying ENSO phases or nutrient availability? 3) How do changes in environmental conditions influence protist community composition and diversity across space and time? Answering these fundamental questions will serve to elucidate a clearer understanding of phytoplankton dynamics in this highly variable and ecologically significant region.

2 | Materials and Methods

2.1 | Sample Collection

Five two-week long research cruises were conducted throughout the Galápagos Marine Reserve (GMR) in October/November of 2014, 2015, 2016, 2018 and 2022. A total of 24 distinct sites were sampled across the years, although not all sites were sampled each year due to logistical constraints. All sites were within an area spanning 1° N to 2° S and 92° W to 89° W with a water column depth of at least 100 m. Stations located within the cove of Isabela (2, 3, 4, 5, 6, 7, 8 and 29) were classified as Equatorial Undercurrent Regions and all other sites were classified as the Eastern Region (Figure 1a) in accordance with Jang et al. (2022).

CTD profiles were obtained at each site and additional discrete samples of seawater were collected for biological and chemical measurements at select sites. CTD casts (SeaBird SBE 19plus) were deployed to 100 m in depth, measuring temperature, salinity, photosynthetic active radiation (PAR) and chlorophyll-*a* (chl-*a*) fluorescence throughout the water column. Following the CTD cast, two separate casts using 10 L Niskin bottles set up

in series were performed to collect discrete water-column samples throughout the euphotic zone at depths corresponding to 50%, 30%, 10% and 1% of the incident irradiance (I_0) based on the previously measured PAR profile from the CTD cast. Seawater was dispensed into acid-cleaned, seawater-rinsed 10 L cubitainers and subsampled for measurements including dissolved inorganic nutrients (nitrate [NO_3^-], phosphate [PO_4^{3-}] and silicic acid [$Si(OH)_4$]), size-fractionated chl-*a*, size-fractionated dissolved inorganic carbon (DIC) and NO_3^- uptake rates, and small phytoplankton cell counts via flow cytometry (*Synechococcus*, *Prochlorococcus* and picoeukaryotes). Additional seawater was collected at the 50% I_0 depth for 18S rRNA gene sequencing analyses.

2.2 | Seawater Properties

Temperature and salinity measurements were used to determine the depth of the mixed and sub-thermocline layers. The CTD casts were corrected using SeaBird's SeaSoft software to account for sensor drift between calibrations. MATLAB (R2017b) was used to calculate potential density, using the sw_pden() function from the Mixing Oceanographic toolbox v1.8.0.0. The surface mixed layer depth was defined as the depth at which the change in density from the surface was greater than 0.35 kg m^{-3} (Holte and Talley 2009). The sub-thermocline layer depth (or the top of the deep layer) was determined by calculating the depth at which changes in density from the bottom of the cast were greater than 0.2 kg m^{-3}.

Dissolved nutrient samples were first filtered through a glass fiber filter (GF/F, nominal porosity 0.7 mm) and the filtrate was stored at −20°C in acid-cleaned plastic scintillation vials. Dissolved nutrient concentrations were measured using the OI Analytical Flow Solutions IV auto-analyser (Parsons et al. 1984) by the Wetlands Biogeochemistry Analytical Services at Louisiana State University. Detection limits were 0.09 μmol L^{-1} for nitrate, 0.02 μmol L^{-1} for phosphate and 0.02 μmol L^{-1} for silicic acid. Reference standards for dissolved nutrients in seawater were also run to ensure quality control.

2.3 | Phytoplankton Biomass and Productivity

Phytoplankton biomass was estimated by measuring size-fractionated (< 5 and ≥ 5 μm) chl-*a* concentration. The samples were obtained from 400 mL of seawater passed through a filter cascade containing a 5 μm polycarbonate filter (47 mm, Millipore) by gravity filtration and a 0.7 μm nominal porosity glass fibre filter (25 mm, Whatman GF/F) under gentle vacuum pressure (< 100 mmHg). Chl-*a* was then extracted from these filters using 90% acetone and fluorescence values were measured using a Turner Designs 10-AU fluorometer according to Parsons et al. (1984). For flow cytometry, 1.8 mL of seawater was placed into a cryovial containing 0.2 mL of paraformaldehyde (PFA, final concentration: 10%) solution according to Marie et al. (2005). Samples were then incubated on ice for 15 min prior to being frozen at −20°C. Subsequently, picophytoplankton cells including *Prochlorococcus*, *Synechococcus* and picoeukaryotes were enumerated using flow cytometry following Johnson et al. (2010) on the BD FACSCalibur (Figure S6).

To measure size-fractionated particulate carbon (POC), particulate nitrogen (PON), DIC uptake rates (DICuptake) and NO_3^- uptake rates (Nuptake), 618 mL subsamples were spiked with $NaH^{13}CO_3$ and $Na^{15}NO_3$ to achieve final concentrations of 120 and 0.5 μmol L^{-1}, respectively. These were incubated for 24 h in open top, transparent plastic flow-through incubators covered with neutral density screening to mimic the light intensity observed at the depth the samples were collected (Slawyk et al. 1977). The addition of 15 N stable isotope was intended to target additions of ~10% ambient concentrations, assuming average nitrate concentrations of 5 μmol L^{-1}, although actual measured nitrate concentrations were used in the uptake calculation. Subsequently, samples were filtered through a filter cascade including a 5 μm polycarbonate filter (47 mm) by gravity filtration and a pre-combusted (450°C for 5 h) GF/F filter (25 mm) by gentle vacuum (< 100 mmHg). Cells captured on the 5 μm polycarbonate filter were rinsed onto a pre-combusted GF/F filter. The filters were kept frozen at −20°C until return to the lab, where they were dried in an oven at 60°C for 24 h. Then, the filters were pelletised using tin squares (Elemental Microanalysis) and quantified using an elemental analyser paired with an isotope ratio mass spectrometer (EA-IRMS) at the UC Davis Stable Isotope Facility. Measurements of PN and PC were obtained simultaneously with uptake rates of NO_3^- and DIC. Dissolved NO_3^- concentrations, PN, PC and ^{15}N and ^{13}C atom percentages were used to calculate volumetric NO_3^- uptake and DIC uptake rates of the different size fractions (Dugdale and Goering 1967). At sites where sampled depths had NO_3^- concentrations below 5 μmol L^{-1}, uptake rates may have been stimulated by isotope addition and thus would indicate potential uptake rate rather than absolute uptake rate.

2.4 | 18S rRNA Gene Sequencing and Bioinformatics

Four litres of seawater from the 50% I_0 depth were filtered using an in-line vacuum (< 100 mmHg) through a 0.45 μm NES membrane filter (47 mm, Pall). Pre-screening was not performed to prevent screening out large chain-forming diatoms or mixotrophic protists. Samples ($n = 85$) were then stored frozen at −20°C till DNA extraction. The V4 region of the 18S rRNA gene was sequenced to obtain taxonomic identification of protistan communities. DNA was extracted from the filters using the DNeasy Plant Mini Kit (QIAGEN). Cells collected on the filter were lyzed with 0.2 g of sterile glass beads and 400 μL of buffer AP1 (QIAGEN) using a bead beater at 48 rpm for 1 min. Subsequent extraction and purification were performed according to the manufacturer's instructions. 18S rRNA gene fragments were amplified via polymerase chain reaction (PCR) using the following V4 primer sets: 18S forward (5'-CAGCASCYGCGGTAATTCC-3′) and reverse (5′-ACTTTCGTTCTTGAT-3′) as described by Lin et al. (2017) and contained 6 bp barcodes on each primer which were used for library multiplexing. PCR reactions were prepared using half reactions of the Ex Taq DNA Polymerase Hot-Start kit (Takara Bio) to obtain a final reaction volume of 25 μL. PCR was conducted using a thermocycler with an initial activation step at 98°C for 1 min, followed by 4 three-step cycles of 94°C for 30 s, 45°C for 45 s, 72°C for 60 s, another 30 three-step cycles of 94°C for 30 s, 57°C for 45 s, 72°C for 60 s and a final

extension step of 72°C for 2 min. PCR was performed in triplicates for each sample and was checked using 1% agarose gel electrophoresis to ensure a clear band was obtained. The PCR products were purified using the QIAquick PCR Purification kit (QIAGEN) and quantified using a Qubit dsDNA assay (ThermoFisher). Amplicons from each sample site were pooled in equimolar amounts and submitted to UNC's High Throughput Sequencing Facility (HTSF) for sequencing on the Illumina MiSeq platform (2 × 300 bp). Samples were processed and sequenced within 2–4 months after cruise completion, targeting a minimum read depth of 10,000 per sample.

Demultiplexing was carried out using tools from QIIME 1.9.1 and Cutadapt 1.18 (Martin 2011). The average read count was ~80,000 reads after demultiplexing. The DADA2 tool (Callahan et al. 2016) in the software pipeline QIIME 2.0 (Bolyen et al. 2019) was used to denoise and join paired ends and remove chimeras. Resultant features were then clustered at a 99% similarity to obtain operational taxonomic units (OTUs). Taxonomic assignment of OTUs were performed using the PR2 database (version 5.0.0, Guillou et al. 2013) and only annotations with a confidence level greater than 0.7 were retained. OTUs classified as Metazoans were excluded from the analysis. Following DADA2 and OTU clustering, 5195 operational taxonomic units (OTUs) were identified. After filtering out non-Eukaryotic taxa, Metazoans and low confidence annotations, 4588 unique OTUs remained that were utilised in downstream diversity and compositional analyses. To maximise data availability, rarefaction was not performed but reads were normalised to relative proportions before subsequent analysis.

2.5 | Data Analysis

Analyses of oceanographic measurements including water properties, phytoplankton biomass and productivity were performed using R version 4.3.2 (R Core Team 2023). Given that DNA samples were only collected at the 50% I_0 depth in conjunction with the fact that most phytoplankton activity is expected to occur within the mixed layer, only chemical and biological measurements taken at the 50% I_0 depth were included in these analyses. A principal component analysis (PCA) was performed using the built-in R function prcomp() to evaluate similarities between environmental metrics at each station. The vegan package (Oksanen et al. 2020) was utilised to calculate alpha (Shannon index) and beta (Bray-Curtis distance) diversity metrics based on the 18S OTU read counts and to perform a non-metric multidimensional scaling (NMDS) analysis to understand variations in composition across the sampled years. Statistical significance was determined using the Wilcoxon rank sum test (rstatix::wilcox_test) when comparing between two independent groups and using PERMANOVA (vegan::adonis2, no. of permutations: 999) when evaluating multivariate data.

Detection of co-occurring phytoplankton taxa was achieved using the WGCNA package in R (Langfelder and Horvath 2008). To account for low counts and many zeros, Hellinger transformation was applied to relative abundance data. The OTUs were inspected to ensure that they were present in at least four samples and did not exhibit zero variance following WGCNA default settings. Modules were detected using the WGCNA::blockwiseModules function with power set to 9 (determined using WGCNA:: pickSoftThreshold function), both networkType and TOMType set to "signed" and minimum module size set to 250. Briefly, this functions to construct a signed network by performing weighted Pearson correlations between all OTUs, considering only positive correlations. WGCNA then identifies modules using unsupervised hierarchical clustering followed by branch cutting using Dynamic Tree Cut (Langfelder et al. 2008). Each module is randomly labelled with a colour and all unassigned OTUs are placed in the grey module. Module eigengenes (MEs) are defined as the first principal component of a given module and can be considered representative of the taxonomic profiles in a module. The MEs are subject to Pearson correlation with key oceanographic measurements collected during each survey to determine how environmental factors shaped the community structure of phytoplankton within the euphotic zone. Module richness was defined as the number of OTUs assigned to each module. Module diversity was calculated by sub-setting the OTUs assigned to each module and applying the Shannon index to the OTU subset for each station. The average diversity across all stations was considered the module diversity for that OTU subset.

3 | Results

In efforts to characterise both spatial and temporal differences in environmental conditions, 24 distinct stations were sampled across the Galápagos archipelago over the span of five years (Figure 1a). The five sampling years included an El Niño event (2015), a La Niña event (2022) and three ENSO neutral years (2014, 2016 and 2018) as defined by the Niño 1 + 2 region sea surface temperature anomalies during our observation periods (Ocean Observations Panel for Climate (OOPC) 2025).

3.1 | Physical Seawater Properties

Temperature and salinity were negatively correlated ($p < 0.01$, Figure S1) across the dataset, likely due to upwelling of EUC waters which are cooler and more saline than surface waters. Mixed layer waters averaged 24.6°C (SD = 0.81°C) under El Niño conditions, 19.7°C (SD = 1.69°C) under La Niña conditions and 22.1°C (SD = 1.05°C) under ENSO neutral conditions (Figure 1b; Figure S2). Measurements taken during CTD casts (Figure 2a) indicate that under El Niño conditions, warmer and more saline conditions were observed near the surface whereas under La Niña conditions, surface waters were noticeably cooler and fresher. The three ENSO neutral years displayed more intermediate levels of temperature and salinity.

Sea surface temperatures were consistently lower within the stations west of Isla Isabela, corresponding to the occurrence of persistent upwelling in this region, bringing up cooler waters from beneath the subsurface waters. Across the five sampling years, mixed layer water temperatures averaged 20.3°C (SD = 2.98°C) within the EUC region and 22.3°C (SD = 2.03°C)

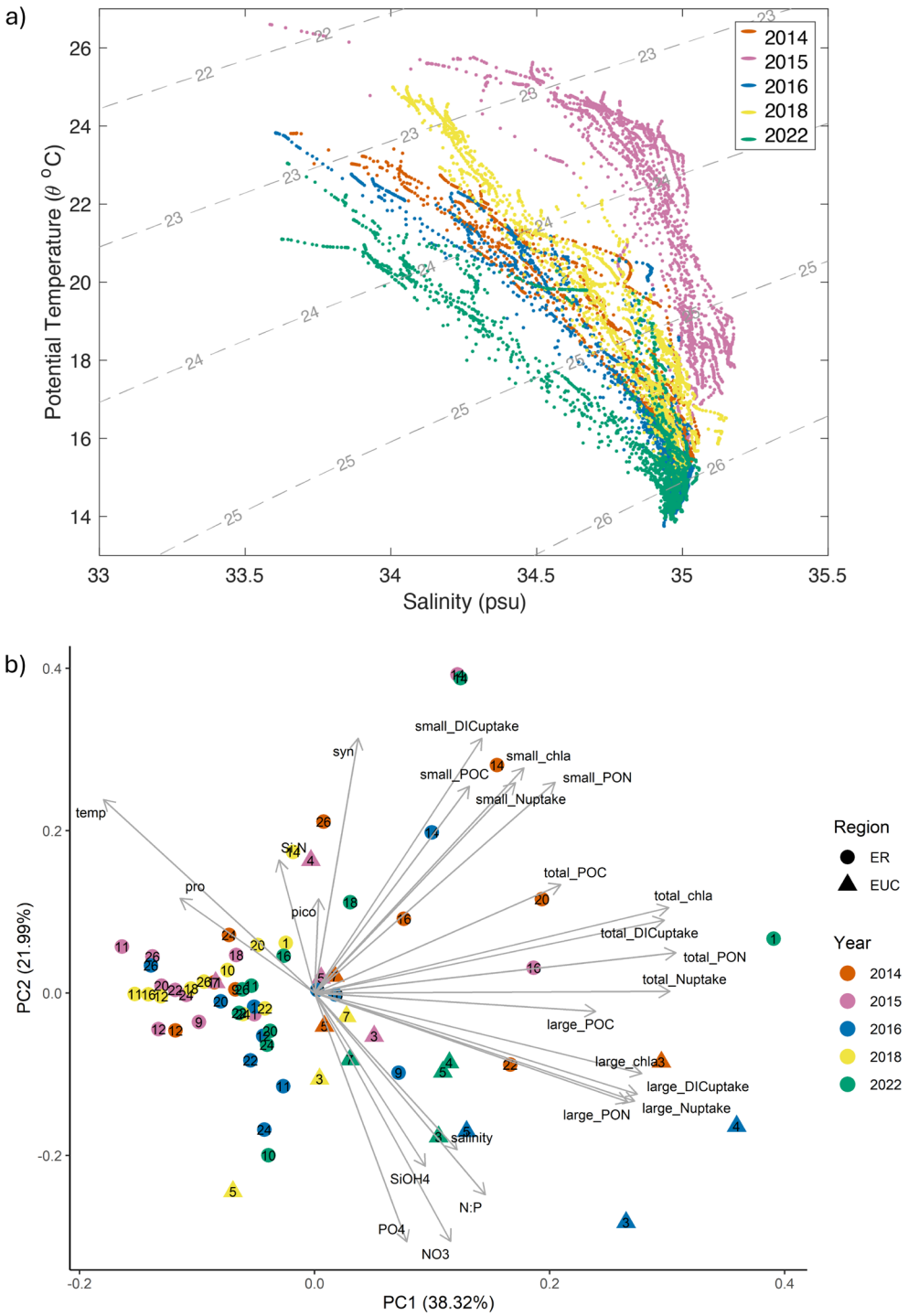

FIGURE 2 | (a) Temperature-salinity profiles obtained during CTD casts to 100 m. Each trajectory represents a single station, with the colours represent sampled year. (b) PCA ordination of oceanographic measurements taken at 50% incident irradiance (depth one) of each station. Points represent a unique station, where the colour represents sampling year, shape represents region and number indicates station number. Grey vectors represent environmental variables including temperature (temp), salinity, nutrient availability (NO_3^-, PO_4^{3-}, $Si(OH)_4$), plankton biomass (chlorophyll-*a*, POC, PON), productivity (DIC & NO_3^- uptake rates), flow cytometry counts (pro = *Prochlorococcus*, syn = *Synechococcus* and pico = picoeukaryotes), nitrate to phosphate ratio (N:P) and silicic acid to nitrate ratio (Si:N). PERMANOVA results indicate statistical significance between both sampling region ($R^2 = 0.011$, $F = 8.12$, $p < 0.001$) and sampling year ($R^2 = 0.15$, $F = 2.77$, $p < 0.001$).

within the ER region, providing evidence that upwelling cooled the mixed layer more in the EUC region. Spatial differences in temperature were significant in 2016, 2018 and 2022 ($p < 0.001$), with the largest difference in temperature being observed in 2016 (Figure 3a). During the El Niño event, no difference in average mixed layer temperatures was observed between the two regions, suggesting a reduced influence of upwelling in the EUC region.

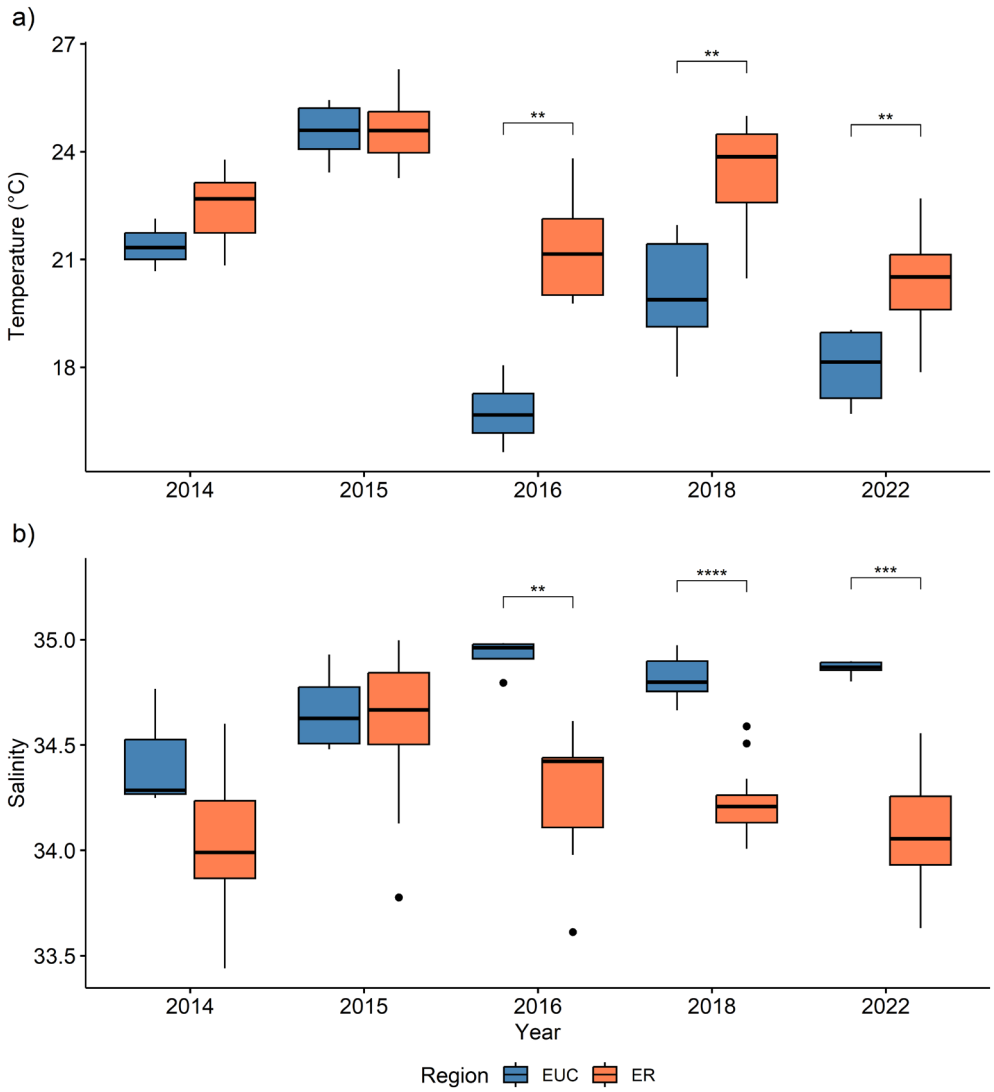

FIGURE 3 | Boxplots displaying mixed layer (a) temperature (°C) and (b) salinity measurements throughout the Galápagos archipelago. Colour represents sampling region: ER is eastern region and EUC is Equatorial Undercurrent region. Statistical significance between regions were determined using the Wilcoxon test and are indicated by asterisks, where **** $\leq 1\times10^{-4}$, *** ≤ 0.001, ** ≤ 0.01, * ≤ 0.05 and no asterisks indicates not significant. Black dots represent outliers, the horizontal lines represent the median and the whiskers represent the 10–90 percentiles.

3.2 | Chemical and Biological Measurements

Macronutrient (NO_3^-, PO_4^{3-}, $Si(OH)_4$) concentrations were strongly negatively correlated with seawater temperature (Figures 2b and S1), indicating that the bulk of the nutrients available to the Galápagos archipelago was supplied by upwelling of cool, nutrient-rich subsurface waters. Likewise, average nitrate concentrations were generally higher in the EUC region (Figure 4a), although a significant difference was only observed in 2018 ($p<1\times10^{-4}$) due to outliers within the ER region. This trend was not seen in 2015, with very similar average nitrate concentrations being observed in both EUC and ER regions, further supporting the role of the Equatorial Undercurrent in supplying nutrients to the archipelago.

All three nutrients displayed similar patterns in variability (Figures 2b and S1). To determine potential nutrient limitation, NO_3^-:PO_4^{3-} (N:P) ratios were calculated and compared to the Redfield ratio (Figure 4b). Across all five sampling years, N:P ratios were substantially lower than the Redfield ratio of 16:1 (Redfield 1934). This indicates that biological activity may potentially become limited by nitrate availability within the Galápagos, particularly considering the strong correlation between chlorophyll concentration and nitrate uptake rates ($p<0.001$; Figure S1). The observed N:P ratios were further decreased in 2015 and 2018 (the second warmest year sampled), suggesting a possible link between nitrate limitation and warmer seawater temperatures. A simple linear regression indicated a significant negative relationship between temperature and N:P ratio ($R^2=0.59$, $F=96.5$, $p<0.001$). In addition to N:P ratios, $Si(OH)_4$:NO_3^- (Si:N) ratios also provide useful information to assess the potential for diatom growth limitation due to their requirements for silica to build cell walls. The calculated Si:N ratios (Figure 4c) exceeded the Brzezinski (1985) proposed ratio of 15:16 in 2015 and fell below it in 2022, suggesting that diatoms may be more susceptible to nitrate limitation under El Niño conditions whereas Si limitation may occur under La Niña conditions.

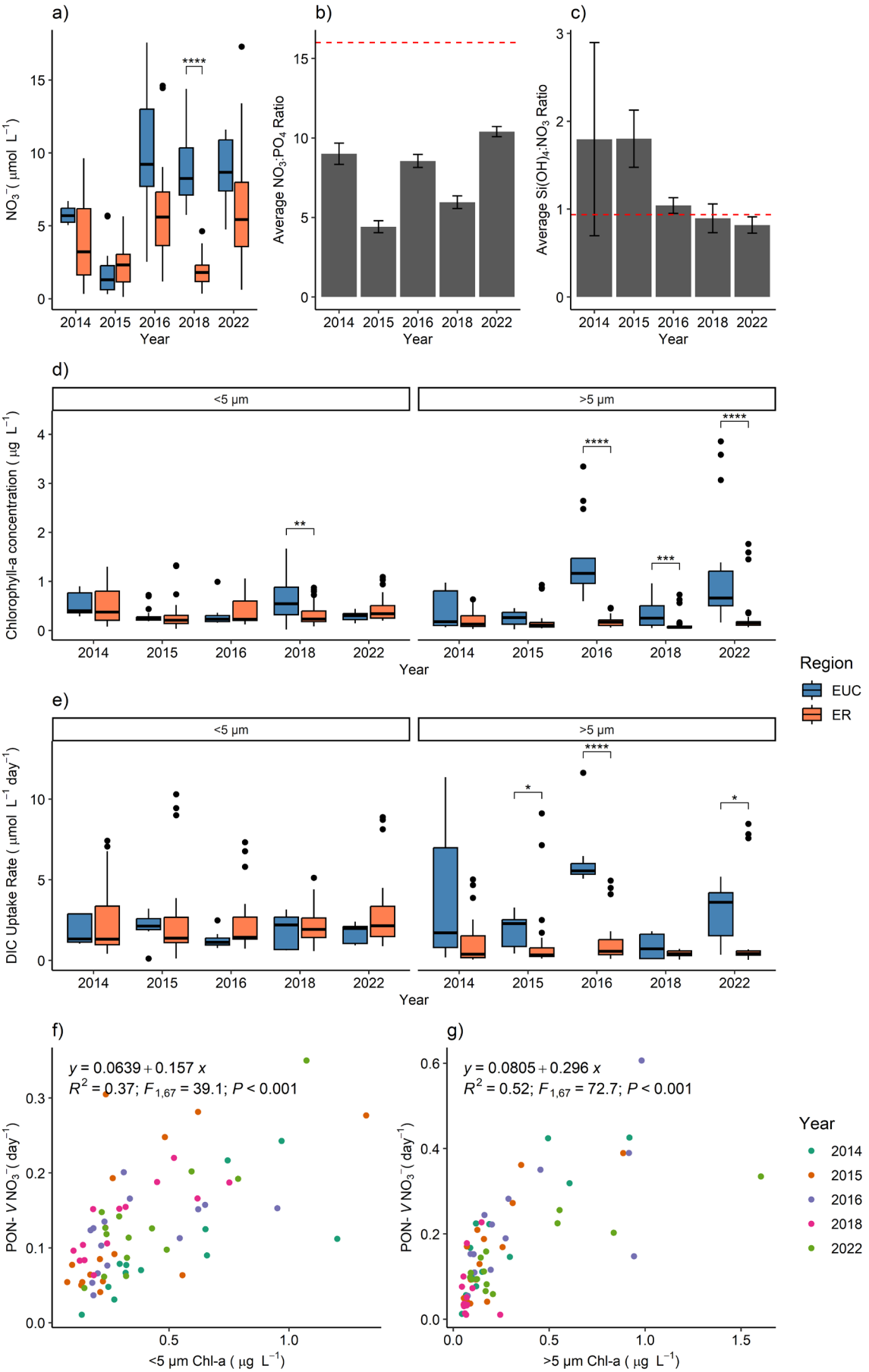

FIGURE 4 | Legend on next page.

FIGURE 4 | Mixed layer chemical and biological measurements throughout the Galápagos archipelago. (a) Boxplots displaying nitrate concentrations ($\mu mol\,L^{-1}$), black dots represent outliers and colour represents sampling region: ER is eastern region and EUC is Equatorial Undercurrent region. (b) Average NO_3^-:PO_4^{3-} ratios for the five sampling years, error bars indicate standard error. The red dotted line indicates the Redfield ratio of N:P = 16:1. (c) Average $Si(OH)_4$: NO_3 ratios for the five sampling years, error bars indicate standard error and the red dotted line indicates the Brzezinski ratio for diatoms of Si:N = 15:16. (d) Distribution of chlorophyll-*a* concentrations ($\mu g\,L^{-1}$) shown by boxplots where black dots represent outliers and colour represents sampling region. Samples were size fractionated to obtain measurements from cells less than 5 μm or equal and greater than 5 μm. (e) Productivity of phytoplankton community measured by DIC uptake rate ($\mu mol\,L^{-1}\,day^{-1}$). Measurements were size fractionated in the same manner as (d). Statistical significance between regions were determined using the Wilcoxon test and are indicated by asterisks, where ****$\leq 1\times10^{-4}$, ***≤ 0.001, **≤ 0.01, *≤ 0.05 and no asterisks indicates not significant. The horizontal lines represent the median and the whiskers represent the 10–90 percentiles. (f, g) Relationship between PON-normalised NO_3 uptake rate (day^{-1}) and size-fractionated chlorophyll-a concentration ($\mu g\,L^{-1}$). Linear regression results displayed on plot; colour represents sampling year.

Biological measurements representing both phytoplankton biomass (chl-*a*, PON and POC) and productivity (DICuptake and Nuptake) displayed similar patterns based on size fraction, with small (< 5 μm) size fraction measurements clustering together and large (> 5 μm) size fraction measurements forming a separate cluster (Figure 2b). Large size fraction measurements were more positively correlated to nutrient concentrations, whereas small size fraction measurements were largely independent of nutrients (Figure S1). Additionally, high biomass and productivity within the small size fraction were strongly associated with ER stations, particularly station 14, as indicated by the length and direction of vectors on Figure 2b. On the other hand, greater biomass and productivity in the large size fraction were more associated with EUC stations, suggesting that the increased bioavailability of nutrients within the EUC region supports larger size fraction phytoplankton.

When comparing chlorophyll-*a* concentrations between ER and EUC regions, large cells appear to be disproportionately affected by spatial variability, whereas small cell chl-*a* concentrations tend to remain similar between the two regions (Figure 4d). Within small cells, a significant difference between regions was observed in 2018 ($p < 0.001$), with slightly elevated biomass observed in the EUC region. Large cells more consistently displayed higher chl-*a* concentrations in the EUC region, particularly in 2016 ($p < 1\times10^{-6}$), 2018 ($p < 1\times10^{-3}$) and 2022 ($p < 1\times10^{-5}$). Large cells seem to be more sensitive to changes in nutrient availability within their environment and rapidly respond to increased nutrients by ramping up their biomass.

Trends in phytoplankton biomass correlate significantly with trends in phytoplankton DIC uptake, a proxy for primary production ($R = 0.68$, $p < 1\times10^{-15}$). Within the small size fraction, phytoplankton productivity remained relatively consistent between the two regions and no significant differences were measured despite greater variability observed in the ER region (Figure 4e). Within the large size fraction, phytoplankton productivity was significantly elevated in the EUC region in 2015 ($p < 0.05$), 2016 ($p < 1\times10^{-4}$) and 2022 ($p < 0.05$). Large cell biomass and productivity were decoupled in 2015 and 2018. In 2015, biomass remained relatively similar between regions while productivity was elevated in the EUC region. In 2018, biomass was elevated in the EUC region, but productivity remained relatively constant.

Another measure of productivity is biomass (i.e., PON)-normalised NO_3^- uptake rates which can be used to infer nitrate assimilation rates. Ambient nitrate concentrations may represent the amount available in the environment but are often a poor predictor of phytoplankton activity due to the lag time between nutrient availability and nutrient uptake (Goldman and Glibert 1983). Nitrate assimilation rates describe the amount of nitrate being incorporated into phytoplankton biomass for growth and can thus be equivalent to a specific growth rate when assuming the majority of PON is algal associated. As expected, both small and large size fractions of phytoplankton demonstrated a positive correlation between nitrate assimilation and chlorophyll-*a* concentration. However, large cells resulted in a greater slope (0.296 compared to 0.157 for small cells; Figure 4f,g), indicating that large cells can more efficiently assimilate nitrate into their biomass, further supporting the observation that the growth rate of large size fraction phytoplankton is more dependent on NO_3 availability.

3.3 | Phytoplankton Community Composition

18S rRNA gene sequencing revealed that Alveolata (includes dinoflagellates), Chlorophyta and Stramenopiles (includes diatoms) were consistently the top divisions detected across the Galápagos archipelago (Figure 5; Figure S7). Including Syndiniales, Alveolata accounted for an average of 43% of reads (25.3% excluding Syndiniales), Chlorophyta accounted for 25.1% of reads and Stramenopiles accounted for 17.7% across all stations and years. Rhizaria was occasionally detected at high proportions, particularly in 2018 at station 26 (98.6%), 2014 at station 26 (33.4%) and 2022 at station 1 (22.8%).

Alpha diversity at each station was calculated using the Shannon Diversity Index (SDI) which accounts for both species richness and evenness, with a larger number indicating greater overall diversity. SDI ranged from 0.21 to 5.92 with a mean of 4.28 (Figure 6a). Interestingly, 2015 exhibited the highest alpha diversity despite having generally lower nutrient availability and phytoplankton biomass. Conversely, stations within the Eastern region tended to have higher alpha diversity compared to stations within the EUC region, although statistical significance was not detected. Beta diversity measures the similarity or dissimilarity between two communities and is visualised using a non-metric multidimensional scaling (NMDS) plot (Figure 6b). This revealed significant differences in community composition across both sampling regions ($p < 0.001$) and sampling years ($p < 0.001$). Samples collected in 2015 formed a tight cluster that aligned with warmer

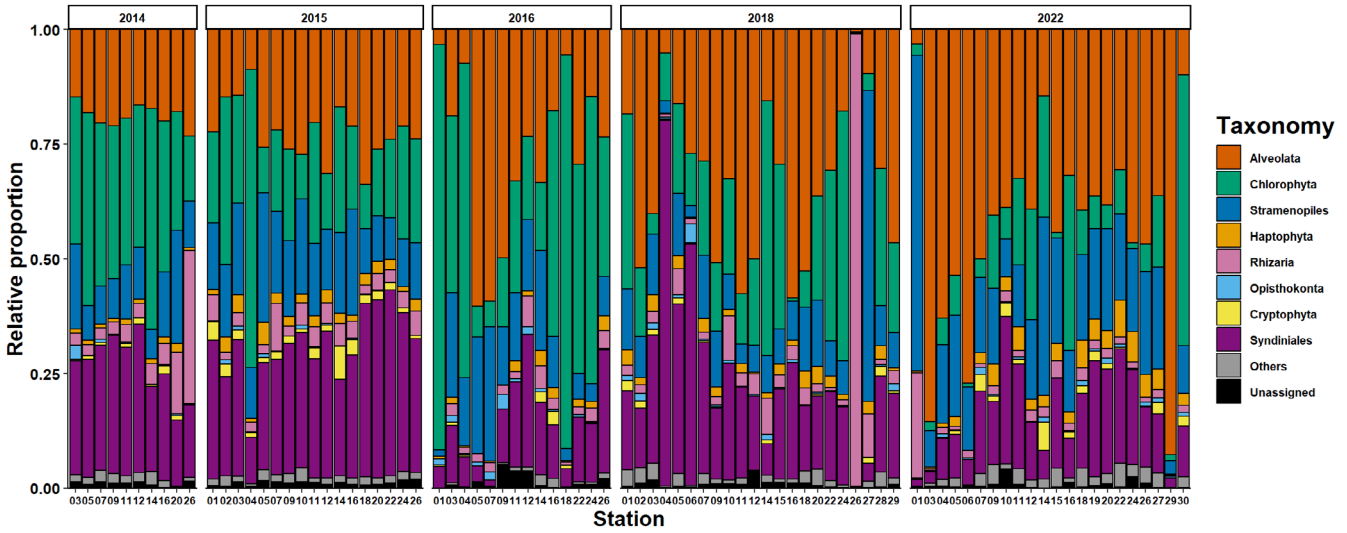

FIGURE 5 | 18S rRNA gene sequencing results annotated to the phyla level according to the PR2 database (version 5.0.0). Read counts were normalised to relative abundance. Syndiniales are part of the Alveolata division but were separated out due to high abundance across most stations.

temperatures, lower nutrient concentrations, small-cell biomass and productivity, and cell concentrations of the cyanobacteria *Synechococcus* and *Prochlorococcus*. The remaining four years had greater variability and were distributed more evenly. Samples collected in the EUC region generally aligned with cooler temperatures, higher nutrient concentrations and large-cell biomass and productivity, although exceptions were noted in all years.

3.4 | Weighted Gene Co-occurrence Network Analysis

In addition to the grey module, four clusters of highly co-occurring OTUs were detected using WGCNA (Figure 7a; Figure S8) representing predicted assemblages of interdependent marine protists. The turquoise module was assigned 406 OTUs, the blue module had 367 OTUs, the brown module had 289 OTUs and the yellow module had 284 OTUs. All remaining OTUs (3242) were placed in the grey module.

The brown module was positively associated with temperature and negatively associated with nutrient concentrations—conditions representative of an El Niño event—as well as *Prochlorococcus* (Figure 7a). The turquoise module exhibited similar trends to the brown module but was not significantly correlated. The blue module was not significantly correlated with any variable apart from a weak negative relationship with salinity. The yellow module was negatively associated with temperature and positively associated with salinity and nutrient concentrations—conditions representative of strong upwelling. The grey module had similar associates to the yellow module and was also positively correlated with small phytoplankton biomass and productivity, *Synechococcus* and picoeukaryote counts.

OTU richness is demonstrated by the number of OTUs assigned to each module and their associated taxonomy (Figure 7b). The four modules contain large proportions of Syndiniales OTUs. The greatest proportions of Alveolata and Stramenopiles OTUs are seen in the turquoise and blue modules. Chlorophyta are most observed in the brown module. To account for overall module composition, alpha diversity was calculated using subsets of OTUs assigned to each module (Figure 7c). The greatest diversity was observed in the blue and brown modules while the yellow module was the least diverse.

To further evaluate the abundance and prevalence of each module, the OTUs in each module were plotted based on their relative proportional read counts across all the sampled stations (Figure 8), depicting which stations and taxa contributed most to each module's composition. Aligning with trends observed in Figure 7a, the brown module was mostly represented in 2015 during the El Niño event and in 2014. This module was heavily dominated by Syndiniales and contained noticeable proportions of Chlorophyta and Rhizaria compared to other modules. The turquoise module was also represented in 2015, particularly in stations 5 and 10, which are located on the western side of the archipelago. Noticeably, Stramenopiles are observed in high proportions in this module. The blue module was primarily detected in 2022 during the La Niña event and represents high proportions of Alveolata, Stramenopiles and Haptophyta. The yellow module was only represented in 2018 at stations 3, 5 and 7, which are all within the EUC region. This module was also dominated by Syndiniales but had elevated proportions of Rhizaria compared to other modules.

4 | Discussion

The Galápagos Islands are renowned for high biodiversity, including many charismatic flora and fauna. This biodiversity is largely supported by the marine ecosystem in the region, fuelled by the upwelling of nutrient-rich waters from the Equatorial Undercurrent. Although the importance of phytoplankton to the Galápagos ecosystem functioning is well understood, few studies have thoroughly characterised the community structure of planktonic protists around the archipelago (e.g., Carnicer et al. 2019; Jang et al. 2022; Neave et al. 2021). This study provides a comprehensive overview of the data collected across five

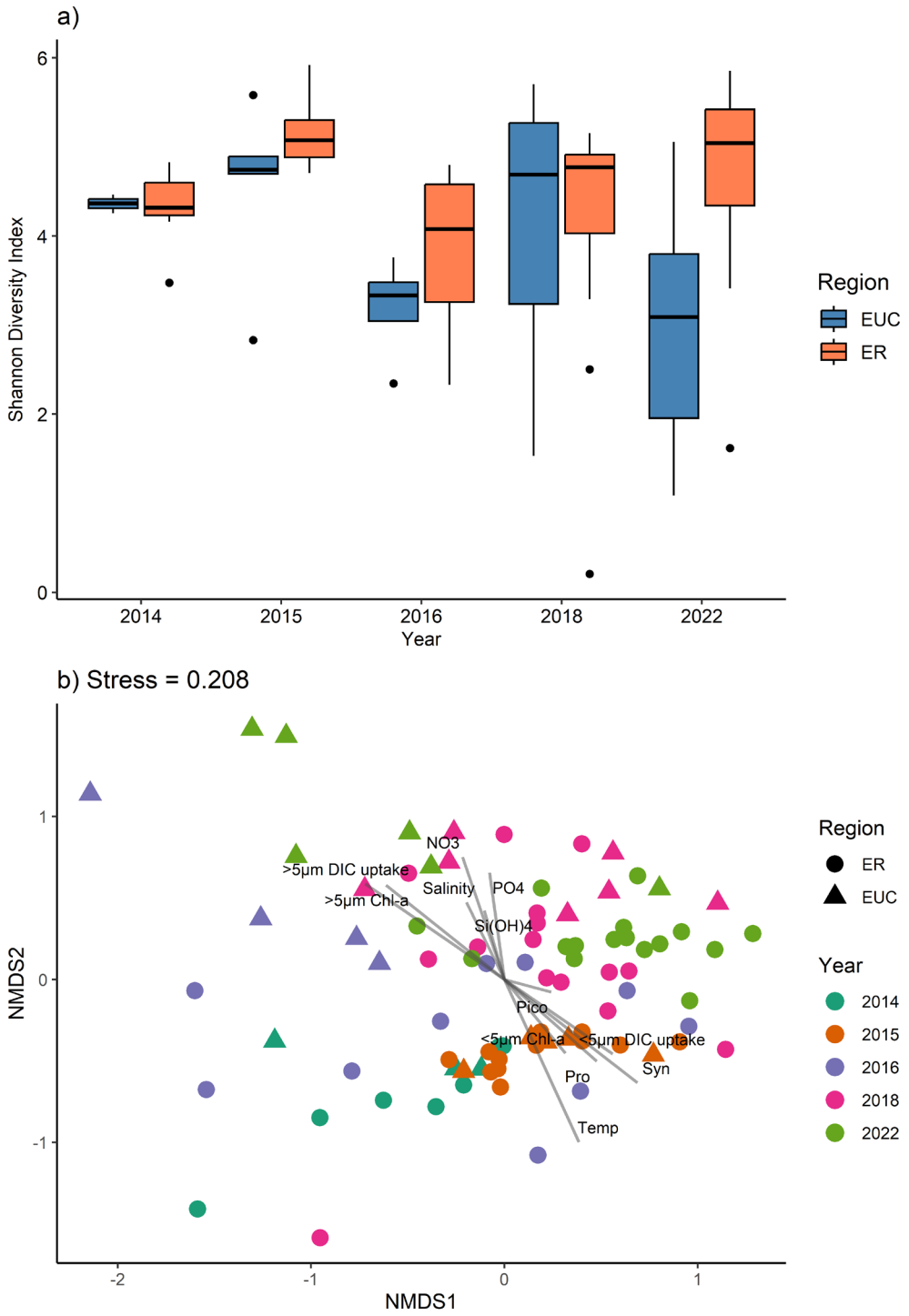

FIGURE 6 | Alpha and beta diversity based on 18S rRNA gene results. (a) Alpha diversity calculated using the Shannon diversity index. Dots indicate outlier points and colour represents sampling region. (b) NMDS ordination plot showing Bray-Curtis distances. Environmental variables are overlayed as grey vectors, shape indicates sampling region and colour represents sampling year. PERMANOVA results indicate statistical significance between both sampling region ($R^2 = 0.04$, $F = 3.41$, $p < 0.001$) and sampling year ($R^2 = 0.22$, $F = 5.54$, $p < 0.001$).

sampling years around the archipelago, spanning both an El Niño (2015) and a La Niña event (2022).

4.1 | Oceanographic Measurements

Nutrient distributions in the mixed layer were strongly influenced by upwelling of cold EUC waters as indicated by the strong negative correlation observed between temperature and nutrient concentrations (Figures 2b and S1). As such, it was unsurprising to measure higher nutrient concentrations (Figures 4a and S3) within the EUC region in four out of the five sampled years. During the El Niño event in 2015, upwelling was severely impeded by the weakening of trade winds and the Equatorial Undercurrent, leading to warmer temperatures across the archipelago in both the EUC and ER

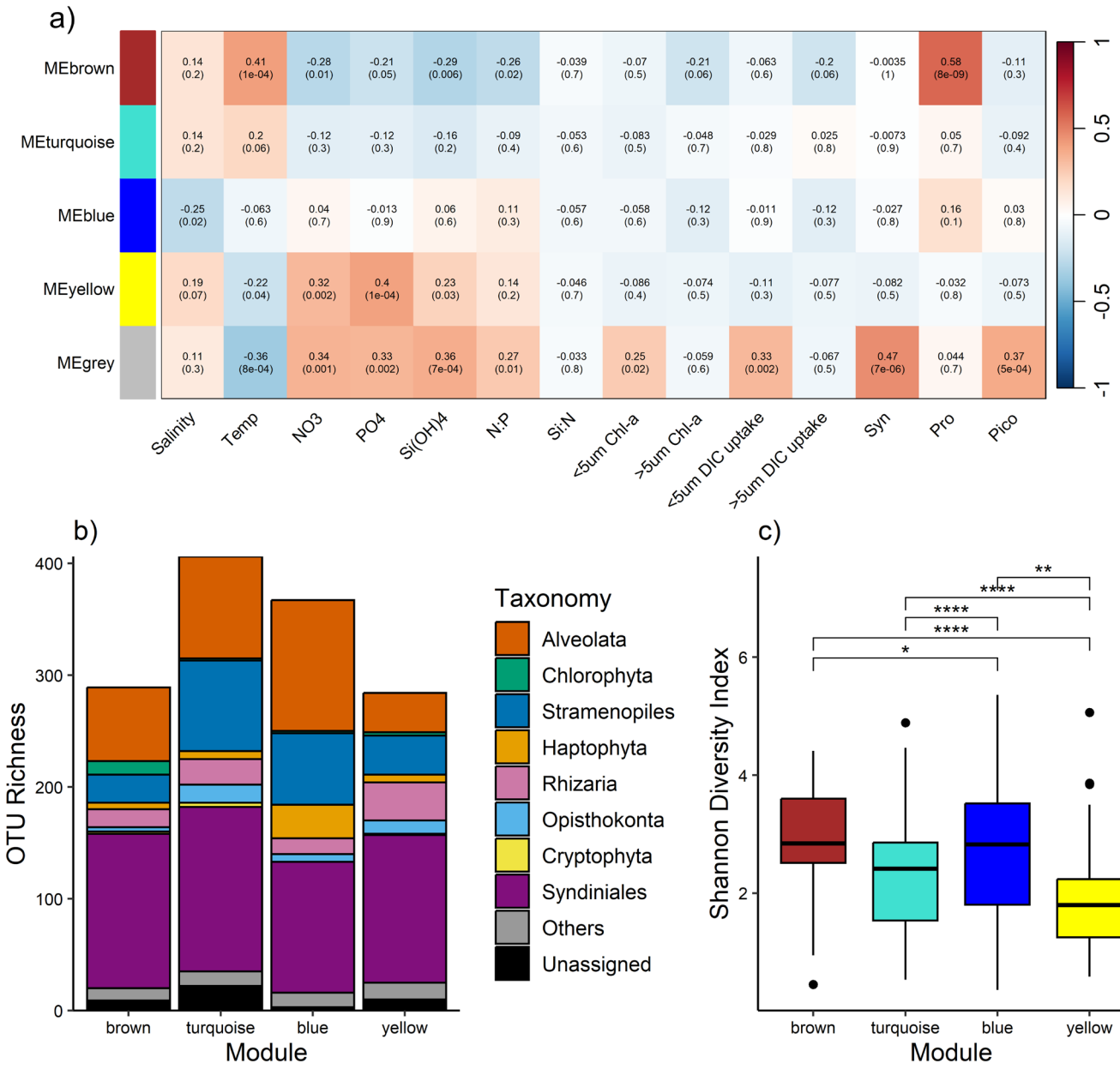

FIGURE 7 | Weighted gene co-occurrence network analysis (WGCNA) results. (a) Correlation matrix showing the relationship between each module eigengenes and key oceanographic measurements. Correlation coefficients are indicated by the top number as well as colour scale bar, *p*-values are indicated by the number within parentheses. (Syn = *Synechococcus*, Pro = *Prochlorococcus*, Pico = picoeukaryotes), (b) Operational taxonomic unit (OTU) richness, i.e., number of OTUs assigned to each module along with phyla level taxonomic annotations, (c) Average Shannon diversity index of each WGCNA module. Statistical significance between regions were determined using the Wilcoxon test and are indicated by asterisks, where ****$\leq 1 \times 10^{-4}$, ***≤ 0.001, **≤ 0.01, *≤ 0.05 and no asterisks indicates not significant.

regions. Consequently, nutrient availability was the lowest in 2015 regardless of region. Conversely, the La Niña event in 2022 strengthened the flow of the Equatorial Undercurrent and thus was expected to cause significantly lower temperatures and elevated nutrient availability. However, this was not entirely the case—although sea surface temperatures in 2022 were cooler, they were generally comparable to the temperatures recorded in 2016, an ENSO neutral year. In fact, temperatures in 2022 exceeded those recorded in 2016 at multiple stations and even occasionally exceeded temperatures recorded in 2018 (Figures 3a and S2). Similarly, nutrient levels were relatively comparable between 2016 and 2022, with the highest nitrate concentration being measured in 2016 at station 3 (Figure S3). A potential explanation for this observation could be increased stratification between the sub-thermocline layer and the mixed layer. While the Equatorial Undercurrent waters were detected throughout the region, transporting cooler waters throughout the archipelago, warm fresher waters likely associated with the Equatorial Front may have potentially created increased stratification that prevented efficient entrainment of these nutrient-rich EUC waters into the euphotic zone. Indeed, environmental conditions within and around the archipelago are influenced by many mechanisms beyond just the EUC, including physical processes affecting

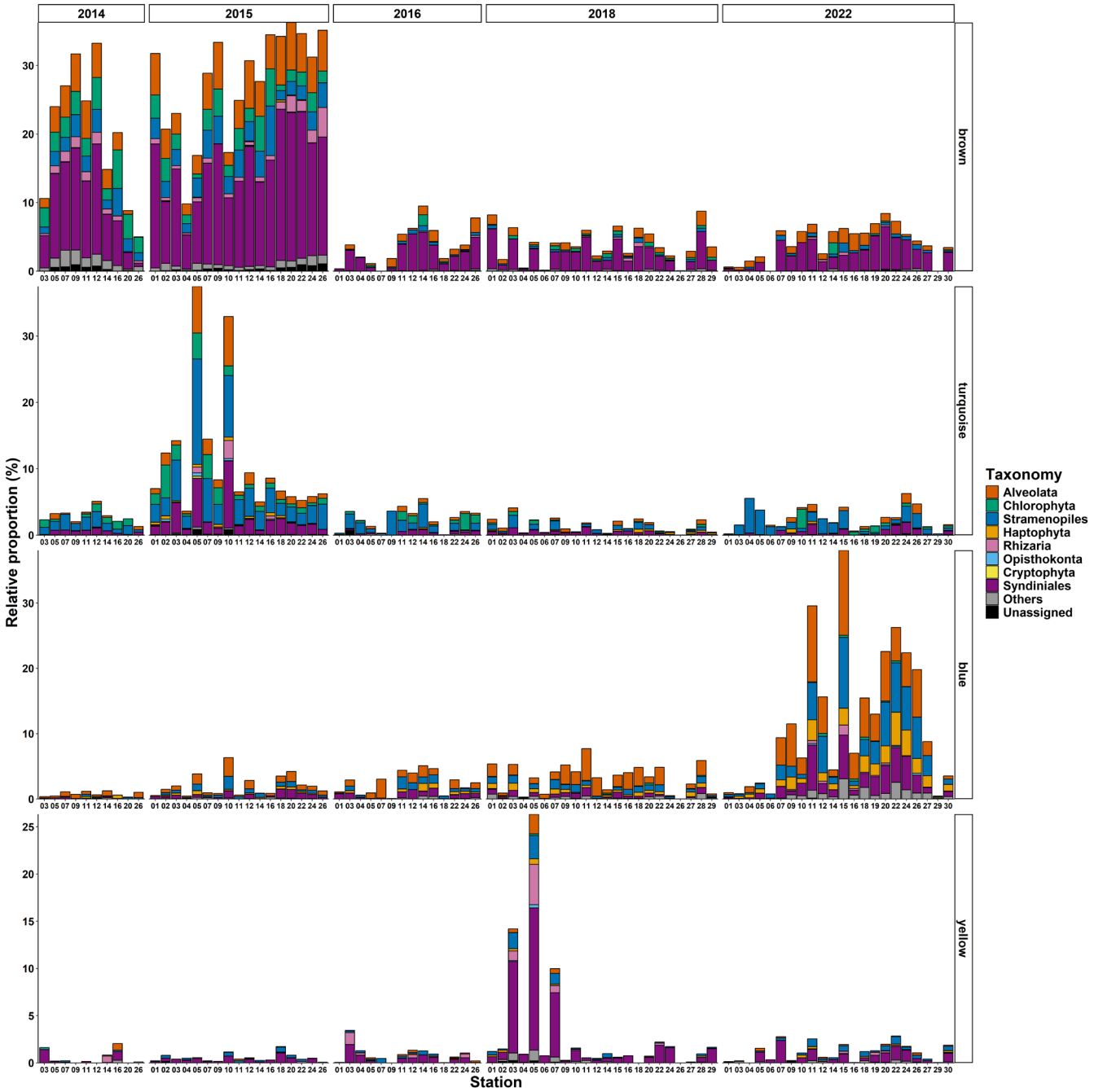

FIGURE 8 | Distribution and relative abundance of the OTUs assigned to each WGCNA module for each station across the five sampling years.

the mixing of deep, nutrient-rich waters into the upper water column, such as internal waves and island wakes (Rudnick et al. 2021; Sweet et al. 2009).

Beyond nutrient availability, nutrient ratios can also provide important information about the functioning of a phytoplankton community. Over the five sampling years, the ratio of nitrate to phosphate was consistently lower than the canonical Redfield ratio of 16:1 (Figure 4b), indicating that phytoplankton productivity could potentially become limited by nitrate availability, particularly considering the strong correlation detected between nitrate uptake rates and chlorophyll-*a* concentration (Figure S1). Similarly, Brzezinski (1985) established that the average silica to nitrogen ratio of diatoms is around 15:16. With some variability, four of the five sampled years had Si:N ratios comparable to the Brzezinski ratio (Figure 4c). A clear exception was observed in 2015 where the Si:N ratio exceeded 15:16, again suggesting that diatoms may be more susceptible to nitrate limitation under El Niño conditions compared to the other essential nutrients.

Stations with higher nutrient availability generally also had higher phytoplankton biomass (Figures 4a,d, S2 and S3). However, nutrient concentrations appeared to correlate more strongly to large cell ($\geq 5\,\mu m$) phytoplankton biomass compared to small cell ($< 5\,\mu m$) groups (Figures 2b and S1). Additionally, PON-normalised NO_3^- uptake rates influenced changes in $\geq 5\,\mu m$ biomass to a greater magnitude than $< 5\,\mu m$

biomass (Figure 4f,g), indicating that larger cells are more efficient at assimilating nitrate into their biomass. In nutrient replete conditions, larger phytoplankton such as diatoms (division Stramenopiles) and dinoflagellates (division Alveolata) often have a competitive advantage over their small-cell counterparts due to higher nutrient uptake rates (Litchman et al. 2007) and storage capacities (Marañón et al. 2013). This explains why larger cells tend to dominate regions with greater nutrient availability and have increased sensitivity in response to nutrient inputs. In contrast, small cells such as green algae (division Chlorophyta) do not appear to be as sensitive to nutrient inputs but instead dominate under low nutrient conditions. These smaller cells have greater surface area-to-volume ratios and decreased diffusion boundary layers which help to maximise nutrient uptake across the cell surface (Irwin et al. 2006).

Phytoplankton biomass and productivity were tightly correlated in both size fractions, as expected since more phytoplankton can uptake greater amounts of nutrients at any given time. Larger cells can uptake more nutrients due to their storage capabilities (Marañón et al. 2013) and thus regions with greater proportions of large phytoplankton generally also have the highest rates of DIC uptake (Figures 4d,e, S4 and S5). An exception to this is station 14, which had consistently more small size-fraction phytoplankton biomass and uptake rates across the five sampling years. This site is anomalous as it is located within Darwin Bay on Isla Genovesa, where a prominent sill restricts seawater exchange with the larger surrounding ocean. Macronutrient concentrations measured at station 14 were relatively lower compared to other stations (Figure S3), creating conditions which favour small size class phytoplankton. The productivity of $< 5\,\mu m$ phytoplankton groups at station 14 was comparable to and even occasionally exceeded DIC uptake rates of larger phytoplankton groups (Figure S5), demonstrating that small phytoplankton can contribute significantly to primary productivity under certain conditions, a trend which has been previously observed by Meyer et al. (2022).

4.2 | Phytoplankton Community Composition

An advantage of using 18S rRNA gene amplicon sequencing over traditional light microscopy or pigment-based taxonomic analysis is its high taxonomic resolution. Based on these data, Alveolata (including dinoflagellates) appear to be the prevalent division across the Galápagos Archipelago (Figure 5), followed by Chlorophyta and Stramenopiles (including diatoms). However, a notable limitation of 18S metabarcoding is that read counts may be skewed by taxa with larger genomes or high 18S rRNA gene copies (Gong and Marchetti 2019; S. Lin 2011). Thus, instead of focusing solely on relative proportions, the distribution of taxonomy across time and space must also be considered. Chlorophytes (which are generally in the small size class) appeared in greater proportions in 2014 and 2015 as well as ER stations in other years. Alveolata and Stramenopiles (which are generally in the large size class) were observed in higher relative proportions in 2016, 2018 and 2022, particularly within stations that were heavily influenced by the EUC. These trends support prior correlations observed between nutrient availability and phytoplankton size class contribution to total biomass and productivity.

Alpha diversity trends indicated that 2015 exhibited the highest diversity and stations within the ER region were generally more diverse than stations within the EUC region (Figure 6a). This suggests that stations with higher phytoplankton biomass and productivity, such as those in the EUC region, are comparatively less diverse than those in the ER region. As such, it is likely that stations which were heavily dominated by large phytoplankton biomass (e.g., diatoms and dinoflagellates) taking advantage of the increased nutrient availability had reduced community diversity through exclusion by competition. The relationship between phytoplankton diversity and productivity tends to be unimodal with diversity peaking at intermediate levels of productivity (Vallina et al. 2014). Hence, the expected decrease in diversity that corresponds with increases in biomass and production has been documented in other studies involving phytoplankton diversity and community structure (Duarte et al. 2006; Napoléon et al. 2014).

Beta diversity was evaluated using Bray-Curtis distances and visualised using NMDS (Figure 6b). Tight clustering of stations is observed in 2015, whereas other sampled years were more evenly dispersed. This indicates that spatial phytoplankton community structure displayed less variability in 2015 compared to the other years. Environmental conditions were stable across the archipelago in 2015, with little to no variation in temperature (Figures 3a and S3) or nutrient availability (Figures 4a and S3) across regions. During the El Niño event, upwelling was significantly suppressed, as indicated by the higher sea surface temperatures observed in 2015. This resulted in more homogenous conditions across the archipelago and likely explained why phytoplankton community structure remained relatively unchanged. In contrast, the other years, especially the cooler years in 2016 and 2022, experienced greater variability in community structure due to heterogeneous environmental conditions caused by the EUC and other water masses influencing the archipelago. The environmental vectors on the NMDS plot align with all previous trends reported—all 2015 stations and other ER stations align with vectors associated with temperature, small phytoplankton biomass and productivity, as well as cyanobacteria and other picoeukaryotes. EUC stations, particularly those of 2016 and 2022, aligned loosely with nutrients, large phytoplankton biomass and productivity vectors. This indicates that warmer stations had lower nutrient availability due to limited upwelling and were dominated by smaller phytoplankton groups, whereas cooler stations had opposite trends; enhanced upwelling provided greater nutrient availability, which supported the dominance of larger phytoplankton groups.

4.3 | Phytoplankton Community Structure

To deduce how phytoplankton community structure was affected by environmental variability, weighted gene co-expression network analysis (WGCNA) was performed to identify clusters (termed modules) of highly correlated OTUs and relate each module to external measurements (Langfelder and Horvath 2008). While initially designed for gene expression data, the WGCNA technique has proved useful in analysing microbial

composition data obtained via amplicon sequencing techniques. Particularly, marine ecologists have been successful in applying WGCNA to phytoplankton composition datasets for identification of community assemblages to understand the influence of environmental factors (e.g., Gong et al. 2020; Y. Lin et al. 2021). WGCNA is uniquely suited to this dataset due to the diverse environmental conditions that exist around the Galápagos Archipelago both spatially and temporally. The intersection of multiple ocean currents brings both warm tropical waters (SEC and NECC) as well as cold nutrient-rich waters (Humboldt current and Equatorial undercurrent) to the archipelago, creating heterogeneous conditions around the region. Additionally, the Galápagos is strongly influenced by the ENSO cycle which shifts every 2–7 years (Litchman et al. 2007; Marañón et al. 2013) from warm El Niño conditions, during which upwelling is heavily suppressed, to cold La Niña conditions, when upwelling is enhanced. As such, it is challenging to broadly categorise stations under different conditions, and utilising unsupervised hierarchical clustering allows for the detection of highly co-occurring modules irrespective of spatial or temporal factors. The modules are subsequently correlated to measured oceanographic metrics to elucidate how these physical, chemical and biological factors interact to shape the phytoplankton communities.

This analysis identified four distinct modules, each representing a diverse community of highly interconnected marine protists. Describing the environmental conditions that promote co-occurrence among specific taxa may offer insight into interactions between marine protists. These modules represented three different environmental conditions: 1) El Niño conditions where modules had a positive correlation with temperature and a negative correlation with nutrients (brown), 2) neutral conditions where modules had weak correlations with nutrient concentrations (turquoise and blue modules) and 3) upwelling conditions where modules had a negative correlation with temperature and a positive correlation with nutrients (yellow module). Each module is treated as a distinct community and their biodiversity is captured through richness and diversity metrics (Figure 7b,c). The turquoise and blue modules had the highest OTU richness, potentially because the neutral environmental conditions associated with these modules allowed for more unique taxa to thrive. Alpha diversity of the yellow module was significantly lower compared to all other modules, likely explained by high nutrient concentrations leading to the dominance of specific taxa that can take advantage of such conditions. Interestingly, alpha diversity of the turquoise module was also relatively low despite it having the highest OTU richness, suggesting an unbalanced community.

Module relative abundance and distribution (Figure 8) provides more insight on each community. Predictably, the brown module was predominantly observed in 2015, during the El Niño event, as well as in 2014. Compared to the other modules, elevated proportions of Chlorophyta and Syndinales are observed in this module. Chlorophytes which are generally smaller in size are likely to thrive under low nutrient conditions, as discussed earlier. The physiology of Syndiniales are less well understood. Syndiniales are a parasitic group that have been observed in all marine environments, known to infect many organisms ranging from other phytoplankton to zooplankton and even fish (Anderson and Harvey 2020; Clarke et al. 2019). As such, Syndiniales are ubiquitous in any 18S rRNA gene amplicon dataset yet remain chronically understudied. The fact that most Syndiniales occur within the brown module during the El Niño event could suggest that Syndiniales may favour such conditions.

The turquoise module was also mostly observed in 2015, particularly stations 5 and 10. This module had noticeably higher proportions of Stramenopiles (including diatoms) and was weakly associated with low nutrient conditions. Given diatoms ability to store nutrients within their vacuoles (Jeong et al. 2010), they can persist in depleted nutrient environments preferentially over other large size class protists such as dinoflagellates. However, should nutrient availability become sufficiently low, Chlorophytes become more advantageous and increase in proportion as observed within the brown module.

The blue module was most neutral in associations with oceanographic measurements, except for a slight negative correlation with salinity. Interestingly, this module is most frequently observed in 2022 and contained greater proportions of Alveolata and Haptophyta compared to the other modules. Despite the La Niña event recorded in 2022, nutrient concentrations were not substantially elevated compared to other ENSO-neutral years. This is most likely attributed to greater stratification between fresher surface waters and high salinity deep waters which prevented upwelled waters from penetrating surface layers efficiently. These neutral nutrient conditions appear to be highly favourable to the well-documented mixotrophic groups Alveolata (Jeong et al. 2010) and Haptophyta (Llopis Monferrer et al. 2022). Protists with mixotrophic ability can take advantage of a wide range of nutrient regimes as they are able to utilise grazing to take up additional nutrients while still being able to generate the majority of their energy via photosynthesis.

The yellow module was primarily detected in stations 3, 5 and 7 in 2018, and was associated with high nutrient availability. This module had large proportions of Syndiniales and Rhizaria compared to other modules. Although Rhizaria consists of primarily heterotrophic organisms, they frequently form symbiotic relationships with photosynthetic algae and are important players in biogeochemical cycling (Llopis Monferrer et al. 2022). Still, it is challenging to explain the elevated proportions of Rhizaria in the yellow module solely based on the module's environmental associations.

The grey module is composed of OTUs that were not successfully clustered into a specific module and thus is not considered a distinct protist community. However, the majority of OTUs (3243 out of 4588, 70.7%) detected across the five sampling years were placed in this module and the module is strongly associated with upwelling conditions such as seen in its negative correlation with temperature and positive correlations with macronutrient concentrations. This module is also correlated with $< 5\,\mu m$ chl-*a* concentrations, DIC uptake rates, and *Synechococcus* and picoeukaryotes cell densities. This indicates that most marine protists observed around the Galápagos islands proliferate under strong upwelling conditions and are not highly co-occurring with one another. Under normal upwelling conditions, high nutrient availability allows for a variety of ecological niches and reduces the likelihood of specific communities becoming dominant.

4.4 | Limitations and Considerations

This study provides a comprehensive examination of phytoplankton in relation to environmental conditions in the Galápagos Islands over the span of five years, including both an El Niño and a La Niña. However, there are several limitations that should be noted. Firstly, only single, discrete measurements were collected at each station, representing a snapshot in time. This does not provide any information on physiological processes that might have a lag time. For example, while it is expected that regions with high nutrient availability will have more biomass and higher primary productivity, this dynamic is not always captured as phytoplankton require time to respond to nutrient inputs and ramp up their metabolic activity. In contrast, it is possible to observe high levels of biomass and productivity in areas where low dissolved nutrient concentrations are detected, due to their depletion by biological activity. To capture these sorts of temporal dynamics, repeated time-series measurements need to be performed. Secondly, sampling was only conducted in the cool season (Oct–Nov) that coincides with when the trade winds and upwelling are stronger. This means that seasonal variability throughout the year is not captured in this dataset and phytoplankton dynamics during the warm season are not described.

There are several considerations regarding the WGCNA. Most significantly, the modules are assigned based on correlation between OTUs. While correlation can often represent biological interactions, it can also be somewhat spurious at times. Thus, while the modules obtained by WGCNA were treated as distinct phytoplankton communities in this study, it should be noted that biological interactions between different groups of phytoplankton cannot be determined by correlation alone. Still, this analysis provides a valuable approach to examining ecological communities in areas with great heterogeneity in environmental conditions.

4.5 | Key Takeaways and Recommendations

Environmental conditions around the Galápagos Islands are affected by many factors, including the intensity of the Equatorial undercurrent and other ocean currents, ENSO cycles and more localised effects such as island wakes and internal tides. The western region of the archipelago is more directly affected by upwelling and thus experiences lower temperatures and higher nutrients compared to the rest of the Galápagos. Under warm El Niño conditions, upwelling is suppressed, giving rise to more consistent conditions characterised by high temperatures and low nutrient availability. Phytoplankton community structure is more similar under these conditions as well, though greater variability is observed under cooler ENSO neutral or La Niña conditions. With increasing global temperatures, it is reasonable to expect sea surface temperatures around the Galápagos to increase with time, particularly with influence from warm currents such as the NECC and SEC. This may result in increased stratification between the mixed layer and deeper ocean waters, preventing cold nutrient-rich waters from reaching the surface even during non-El Niño years. Should that happen, the phytoplankton may shift to look more similar to the brown module, with increased prominence of small phytoplankton groups such as Chlorophyta. Similar trends have been projected at the global scale (Rousseaux and Gregg 2015). Small phytoplankton have traditionally been understudied compared to their larger counterparts such as diatoms and dinoflagellates. This is mainly because small phytoplankton are harder to identify using traditional microscopy methods and to isolate for culturing experiments. Still, several studies have demonstrated the importance of small phytoplankton in net community production and carbon export (Juranek et al. 2020; Meyer et al. 2022), emphasising the need for more attention to be placed on these frequently overlooked taxa. Making accurate predictions of phytoplankton compositional changes based on environmental conditions remains challenging due to many nuanced variables that can lead to significantly different community compositions. Long-term monitoring of environmental variables and phytoplankton community across seasonal and climatological time scales (e.g., ENSO) is necessary to fill in these gaps and provide a clearer understanding of phytoplankton ecology in and around the Galápagos Archipelago.

Author Contributions

P.L. led the writing of the original draft, performed visualization, formal analysis, investigation, data curation, and software development. H.S. contributed to software development, methodology, investigation, conceptualization, funding acquisition, and writing – review and editing. O.T. contributed to investigation, methodology, software, and writing – review and editing. E.N. participated in investigation, methodology, and writing – review and editing. S.H.J. contributed to investigation and writing – review and editing. Z.J. contributed to investigation, methodology, and writing – review and editing. S.H. was involved in investigation, software development, and writing – review and editing. S.G. contributed to investigation, conceptualization, and writing – review and editing. N.C. contributed to investigation and writing – review and editing. C.M.M. was involved in investigation and writing – review and editing. M.L. contributed to visualization and writing – review and editing. C.V.P. participated in investigation and writing – review and editing. A.M. led the conceptualization, investigation, funding acquisition, methodology, validation, project administration, resources, supervision, and writing – review and editing. All authors reviewed and approved the final version of the manuscript.

Acknowledgements

We thank scientists and staff from the Galápagos Science Center (GSC), Universidad San Francisco de Quito (USFQ) and Galápagos National Park (GNP) for their logistical support during the cruises. In particular, Diego Páez-Rosas (USFQ), Juan Pablo Muñoz (USFQ), Daniela Alarcon (USFQ), Diana Pazmiño (USFQ), Eduardo Espinoza (GNP), Jennifer Suarez (GNP) and the crew of the M/V Sierra Negra. We are grateful to Steve Walsh (UNC), Carlos Mena (USFQ) and Amanda Thompson (UNC) for their efforts in financially supporting and coordinating the Galápagos Marine Expeditions. Funding for this project was provided to A.M. and H.S. from the Center for Galápagos Studies (Office of the Vice Chancellor for Research), UNC College of Arts and Sciences, and the National Science Foundation grant DEB2326027 and to A.M. from the National Science Foundation grant OCE1751805.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Raw oceanographic measurements collected in this study are available in [Supplemental Data File S1]. Demultiplexed and quality controlled

sequencing data obtained via 18S rRNA gene sequencing were deposited in the SRA under BioProject PRJNA1105929. The processed operational taxonomic units (OTUs) used for taxonomic analysis in this study have been deposited at DDBJ/ENA/GenBank under the accession KIGJ00000000. The version described in this paper is the first version, KIGJ02000000. Additionally, details regarding OTU read counts detected at each station and taxonomic annotations are available in [Supplemental Data File S2] and [Supplemental Data File S3] respectively.

References

Anderson, S. R., and E. L. Harvey. 2020. "Temporal Variability and Ecological Interactions of Parasitic Marine Syndiniales in Coastal Protist Communities." *MSphere* 5, no. 3: 00209-20. https://doi.org/10.1128/msphere.00209-20.

Bolyen, E., J. R. Rideout, M. R. Dillon, et al. 2019. "Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2." *Nature Biotechnology* 37, no. 8: 852–857. https://doi.org/10.1038/s41587-019-0209-9.

Brzezinski, M. A. 1985. "The Si:C:N Ratio of Marine Diatoms: Interspecific Variability and the Effect of Some Environmental Variables." *Journal of Phycology* 21, no. 3: 347–357. https://doi.org/10.1111/j.0022-3646.1985.00347.x.

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." *Nature Methods* 13, no. 7: 581–583. https://doi.org/10.1038/nmeth.3869.

Carnicer, O., P. De La Fuente, A. Canepa, et al. 2019. "Marine Dinoflagellate Assemblage in the Galápagos Marine Reserve." *Frontiers in Marine Science* 6: 00235. https://doi.org/10.3389/fmars.2019.00235.

Chavez, F. P., and R. C. Brusca. 1991. "The Galápagos Islands and Their Relation to Oceanographic Processes in the Tropical Pacific." In *Galápagos Marine Invertebrates. Topics in Geobiology*, edited by M. J. James, vol. 8. Springer. https://doi.org/10.1007/978-1-4899-0646-5_2.

Clarke, L. J., S. Bestley, A. Bissett, and B. E. Deagle. 2019. "A Globally Distributed Syndiniales Parasite Dominates the Southern Ocean Micro-Eukaryote Community Near the Sea-Ice Edge." *ISME Journal* 13, no. 3: 734–737. https://doi.org/10.1038/s41396-018-0306-7.

De Vargas, C., S. Audic, N. Henry, et al. 2015. "Eukaryotic Plankton Diversity in the Sunlit Ocean." *Science* 348, no. 6237: 1261605. https://doi.org/10.1126/science.1261605.

Duarte, P., M. F. Macedo, and L. C. Da Fonseca. 2006. "The Relationship Between Phytoplankton Diversity and Community Function in a Coastal Lagoon." *Hydrobiologia* 555, no. 1: 3–18. https://doi.org/10.1007/s10750-005-1101-9.

Dugdale, R. C., and J. J. Goering. 1967. "Uptake of New and Regenerated Forms of Nitrogen In Primary Productivity1." *Limnology and Oceanography* 12, no. 2: 196–206. https://doi.org/10.4319/lo.1967.12.2.0196.

Ebenezer, V., L. K. Medlin, and J. S. Ki. 2012. "Molecular Detection, Quantification, and Diversity Evaluation of Microalgae." *Marine Biotechnology* 14, no. 2: 129–142. https://doi.org/10.1007/s10126-011-9427-y.

Fiedler, P. C., and L. D. Talley. 2006. "Hydrography of the Eastern Tropical Pacific: A Review." *Progress in Oceanography* 69, no. 2–4: 143–180. https://doi.org/10.1016/j.pocean.2006.03.008.

Goldman, J. C., and P. M. Glibert. 1983. "Kinetics of Inorganic Nitrogen Uptake by Phytoplankton." In *Nitrogen in the Marine Environment*, edited by E. J. Carpenter and D. G. Capone. Academic Press. https://doi.org/10.1016/b978-0-12-160280-2.50015-8.

Gong, W., N. Hall, H. Paerl, and A. Marchetti. 2020. "Phytoplankton Composition in a Eutrophic Estuary: Comparison of Multiple Taxonomic Approaches and Influence of Environmental Factors." *Environmental Microbiology* 22, no. 11: 4718–4731. https://doi.org/10.1111/1462-2920.15221.

Gong, W., and A. Marchetti. 2019. "Estimation of 18S Gene Copy Number in Marine Eukaryotic Plankton Using a Next-Generation Sequencing Approach." *Frontiers in Marine Science* 6: 219.

Guillou, L., D. Bachar, S. Audic, et al. 2013. "The Protist Ribosomal Reference Database (PR2): A Catalog of Unicellular Eukaryote Small Sub-Unit rRNA Sequences With Curated Taxonomy." *Nucleic Acids Research* 41, no. D1: D597–D604. https://doi.org/10.1093/nar/gks1160.

Holte, J., and L. Talley. 2009. "A New Algorithm for Finding Mixed Layer Depths With Applications to Argo Data and Subantarctic Mode Water Formation." *Journal of Atmospheric and Oceanic Technology* 26, no. 9: 1920–1939. https://doi.org/10.1175/2009JTECHO543.1.

Irwin, A. J., Z. V. Finkel, O. M. E. Schofield, and P. G. Falkowski. 2006. "Scaling-Up From Nutrient Physiology to the Size-Structure of Phytoplankton Communities." *Journal of Plankton Research* 28, no. 5: 459–471. https://doi.org/10.1093/plankt/fbi148.

Jang, S. H., P. Lim, O. Torano, E. F. Neave, H. Seim, and A. Marchetti. 2022. "Protistan Communities Within the Galápagos Archipelago With an Emphasis on Micrograzers." *Frontiers in Marine Science* 9: 811979. https://doi.org/10.3389/fmars.2022.811979.

Jeong, H. J., Y. du Yoo, J. S. Kim, K. A. Seong, N. S. Kang, and T. H. Kim. 2010. "Growth, Feeding and Ecological Roles of the Mixotrophic and Heterotrophic Dinoflagellates in Marine Planktonic Food Webs." *Ocean Science Journal* 45, no. 2: 65–91. https://doi.org/10.1007/s12601-010-0007-2.

Jimenez, R. 1981. "Composition and Distribution of Phytoplankton in the Upwelling System of the Galapagos Islands." In *Coastal and Estuarine Sciences*, edited by F. A. Richards, vol. 1, Coastal Upwelling. American Geophysical Union. https://doi.org/10.1029/CO001p0327.

Johnson, Z. I., R. Shyam, A. E. Ritchie, et al. 2010. "The Effect of Iron-and Light-Limitation on Phytoplankton Communities of Deep Chlorophyll Maxima of the Western Pacific Ocean." *Journal of Marine Research* 68, no. 2: 283–308. https://doi.org/10.1357/002224010793721433.

Juranek, L. W., A. E. White, M. Dugenne, et al. 2020. "The Importance of the Phytoplankton "Middle Class" to Ocean Net Community Production." *Global Biogeochemical Cycles* 34, no. 12: e2020GB006702. https://doi.org/10.1029/2020GB006702.

Karnauskas, K. B. 2022. "Whither Warming in the Galápagos?" *PLOS Climate* 1, no. 9: e0000056. https://doi.org/10.1371/journal.pclm.0000056.

Kessler, W. S. 2006. "The Circulation of the Eastern Tropical Pacific: A Review." *Progress in Oceanography* 69, no. 2–4: 181–217. https://doi.org/10.1016/j.pocean.2006.03.009.

Langfelder, P., and S. Horvath. 2008. "WGCNA: An R Package for Weighted Correlation Network Analysis." *BMC Bioinformatics* 9: 559. https://doi.org/10.1186/1471-2105-9-559.

Langfelder, P., B. Zhang, and S. Horvath. 2008. "Defining Clusters From a Hierarchical Cluster Tree: The Dynamic Tree Cut Package for R." *Bioinformatics* 24, no. 5: 719–720. https://doi.org/10.1093/bioinformatics/btm563.

Lin, S. 2011. "Genomic Understanding of Dinoflagellates." *Research in Microbiology* 162, no. 6: 551–569. https://doi.org/10.1016/j.resmic.2011.04.006.

Lin, Y., N. Cassar, A. Marchetti, C. Moreno, H. Ducklow, and Z. Li. 2017. "Specific Eukaryotic Plankton Are Good Predictors of Net Community Production in the Western Antarctic Peninsula." *Scientific Reports* 7, no. 1: 14845. https://doi.org/10.1038/s41598-017-14109-1.

Lin, Y., C. Moreno, A. Marchetti, et al. 2021. "Decline in Plankton Diversity and Carbon Flux With Reduced Sea Ice Extent Along the

Western Antarctic Peninsula." *Nature Communications* 12, no. 1: 4948. https://doi.org/10.1038/s41467-021-25235-w.

Lindley, S. T., and R. T. Barber. 1998. "Phytoplankton Response to Natural and Experimental Iron Addition." *Deep-Sea Research Part II: Topical Studies in Oceanography* 45, no. 6: 1135–1150. https://doi.org/10.1016/S0967-0645(98)00014-9.

Litchman, E., C. A. Klausmeier, O. M. Schofield, and P. G. Falkowski. 2007. "The Role of Functional Traits and Trade-Offs in Structuring Phytoplankton Communities: Scaling From Cellular to Ecosystem Level." *Ecology Letters* 10, no. 12: 1170–1181. https://doi.org/10.1111/j.1461-0248.2007.01117.x.

Llopis Monferrer, N., T. Biard, M. M. Sandin, et al. 2022. "Siliceous Rhizaria Abundances and Diversity in the Mediterranean Sea Assessed by Combined Imaging and Metabarcoding Approaches." *Frontiers in Marine Science* 9: 895995. https://doi.org/10.3389/fmars.2022.895995.

Marañón, E., P. Cermeño, D. C. López-Sandoval, et al. 2013. "Unimodal Size Scaling of Phytoplankton Growth and the Size Dependence of Nutrient Uptake and Use." *Ecology Letters* 16, no. 3: 371–379. https://doi.org/10.1111/ele.12052.

Marie, D., N. Simon, and D. Vaulot. 2005. "Phytoplankton Cell Counting by Flow Cytometry." In *Algal Culturing Techniques*, edited by R. A. Anderson. Elsevier Academic Press. https://doi.org/10.1016/b978-012088426-1/50018-4.

Marin Jarrin, M. J., and T. C. Lippmann. 2019. "Interannual Variability of Mixed Layer Dynamics in the Ecuadorian Ocean." *Journal of Geophysical Research: Oceans* 124, no. 12: 8777–8797. https://doi.org/10.1029/2019JC015086.

Martin, M. 2011. "Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads." *EMBnet.Journal* 17, no. 1: 10. https://doi.org/10.14806/ej.17.1.200.

McCulloch, A. 2011. *A Spatio-Temporal Context for the Phytoplankton Community Patterns of the Galapagos Archipelago and the Northwest Florida Shelf*. North Carolina State University dissertation repository.

Meyer, M. G., W. Gong, S. M. Kafrissen, et al. 2022. "Phytoplankton Size-Class Contributions to New and Regenerated Production During the EXPORTS Northeast Pacific Ocean Field Deployment." *Elementa* 10, no. 1: 00068. https://doi.org/10.1525/elementa.2021.00068.

Napoléon, C., L. Fiant, V. Raimbault, P. Riou, and P. Claquin. 2014. "Dynamics of Phytoplankton Diversity Structure and Primary Productivity in the English Channel." *Marine Ecology Progress Series* 505: 49–64. https://doi.org/10.3354/meps10772.

Neave, E. F., H. Seim, S. M. Gifford, et al. 2021. "Protistan Plankton Communities in the Galápagos Archipelago Respond to Changes in Deep Water Masses Resulting From the 2015/16 El Niño." *Environmental Microbiology* 24, no. 4: 1746–1759. https://doi.org/10.1111/1462-2920.15863.

Ocean Observations Panel for Climate (OOPC). 2025. "Niño1+2." https://stateoftheocean.osmc.noaa.gov/sur/pac/nino12.php.

Oksanen, J., P. Legendre, B. O'Hara, M. H. H. Stevens, M. J. Oksanen, and M. Suggests. 2020. "Vegan: Community Ecology Package. R Package Version 2.5-7." *Community Ecology Package* 10: 3.

Parsons, T. R., Y. Maita, and C. M. Lalli. 1984. *A Manual of Chemical & Biological Methods for Seawater Analysis*. Pergamon Press. https://doi.org/10.1016/c2009-0-07774-5.

Pennington, J. T., K. L. Mahoney, V. S. Kuwahara, D. D. Kolber, R. Calienes, and F. P. Chavez. 2006. "Primary Production in the Eastern Tropical Pacific: A Review." *Progress in Oceanography* 69, no. 2–4: 285–317. https://doi.org/10.1016/j.pocean.2006.03.012.

Pessarrodona, A., J. Assis, K. Filbee-Dexter, et al. 2022. "Global Seaweed Productivity." *Science Advances* 8, no. 37: eabn2465. https://doi.org/10.1126/sciadv.abn2465.

R Core Team. 2023. "R: A Language and Environment for Statistical Computing". R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Redfield, A. C. 1934. "On the Proportions of Organic Derivations in Sea Water and Their Relation to the Composition of Plankton." In *James Johnstone Memorial Volume*, edited by R. J. Daniel. University Press of Liverpool.

Rousseaux, C. S., and W. W. Gregg. 2015. "Recent Decadal Trends in Global Phytoplankton Composition." *Global Biogeochemical Cycles* 29, no. 10: 1674–1688. https://doi.org/10.1002/2015GB005139.

Rudnick, D. L., W. B. Owens, T. M. S. Johnston, K. B. Karnauskas, J. Jakoboski, and R. E. Todd. 2021. "The Equatorial Current System West of the Galápagos Islands During the 2014-16 el niño as Observed by Underwater Gliders." *Journal of Physical Oceanography* 51, no. 1: 3–17. https://doi.org/10.1175/JPO-D-20-0064.1.

Sakamoto, C. M., F. J. Millero, W. Yao, G. E. Friederich, and F. P. Chavez. 1998. "Surface Seawater Distributions of Inorganic Carbon and Nutrients Around the Galapagos Islands: Results From the PlumEx Experiment Using Automated Chemical Mapping." *Deep-Sea Research Part II: Topical Studies in Oceanography* 45, no. 6: 1055–1071. https://doi.org/10.1016/S0967-0645(98)00013-7.

Slawyk, G., Y. Collos, and J.-C. Auclair. 1977. "The Use of the 13C and 15N Isotopes for the Simultaneous Measurement of Carbon and Nitrogen Turnover Rates in Marine Phytoplankton." *Limnology and Oceanography* 22, no. 5: 925–932. https://doi.org/10.4319/lo.1977.22.5.0925.

Sweet, W. V., J. M. Morrison, Y. Liu, et al. 2009. "Tropical Instability Wave Interactions Within the Galápagos Archipelago." *Deep-Sea Research Part I: Oceanographic Research Papers* 56, no. 8: 1217–1229. https://doi.org/10.1016/j.dsr.2009.02.005.

US NASA Jet Propulsion Laboratory Physical Oceanography Distributed Active Archive Center (JPL PO.DAAC). 2011. "GHRSST Level 4 MUR Global Foundation Sea Surface Temperature Analysis (v4.1)." NOAA National Centers for Environmental Information. https://www.ncei.noaa.gov/archive/accession/GHRSST-MUR-JPL-L4-GLOB.

Vallina, S. M., M. J. Follows, S. Dutkiewicz, J. M. Montoya, P. Cermeno, and M. Loreau. 2014. "Global Relationship Between Phytoplankton Diversity and Productivity in the Ocean." *Nature Communications* 5: 4299. https://doi.org/10.1038/ncomms5299.

Worden, A. Z., M. J. Follows, S. J. Giovannoni, S. Wilken, A. E. Zimmerman, and P. J. Keeling. 2015. "Rethinking the Marine Carbon Cycle: Factoring in the Multifarious Lifestyles of Microbes." *Science* 347, no. 6223: 1257594. https://doi.org/10.1126/science.1257594.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.