

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Environmental Genomic Investigations of Phytoplankton in the California Current Ecosystem

Permalink

<https://escholarship.org/uc/item/3wx077tt>

Author

Lampe, Robert

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Environmental Genomic Investigations of Phytoplankton in the California Current Ecosystem

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Oceanography

by

Robert H. Lampe

Committee in charge:

Professor Andrew E. Allen, Chair
Professor Katherine A. Barbeau
Professor Andrew D. Barton
Professor Karsten Zengler

2024

Copyright

Robert H. Lampe, 2024

All rights reserved.

The Dissertation of Robert H. Lampe is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

To all past, present, and future diatomists at Scripps Institution of Oceanography.

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE	iii
DEDICATION	iv
TABLE OF CONTENTS.....	v
LIST OF FIGURES	vi
LIST OF TABLES	xi
ACKNOWLEDGEMENTS	xii
VITA.....	xv
ABSTRACT OF THE DISSERTATION	xviii
Chapter 1 - Molecular Mecanisms for Iron Uptake and Homeostasis in Marine Eukaryotic Phytoplankton	1
Chapter 2 - Short-term Acidification Promotes Diverse Iron Acquisition and Conservation Mechanisms in Upwelling-associated Phytoplankton	41
Chapter 2 Supplemental Information.....	109
Chapter 3 - Relationships between phytoplankton pigments and DNA- or RNA-based abundances support ecological applications	147
Chapter 3 Supplemental Information.....	199
Chapter 4 - Coastal upwelling dynamics govern distinct patterns in diatom abundances, diversity, and gene expression	214
Chapter 4 Supplemental Information.....	267

LIST OF FIGURES

Figure 1.1: Cell surface view of iron uptake and processes that influence iron bioavailability in seawater.....	5
Figure 1.2: (a) Intracellular view of iron transport and storage as well as enzymes associated with iron demand. (b) Iron requirements for photosynthetic electron transport. (c) Iron requirements for mitochondrial electron transport.	9
Figure 2.1: Overview of the initial conditions for the four experiments (E1-E4).	48
Figure 2.2: Physiological measurements associated with the incubations.	52
Figure 2.3: Microbial community composition and diversity for each experiment (E1-4) from both 18S and 16S rRNA amplicons.	57
Figure 2.4: Molecular-level responses to increased pCO ₂	61
Figure 2.5: Potential diatom-related cellular changes from increased acidification driving changes in iron chemistry and/or bioavailability.	71
Supplemental Figure 2.1: Temperature (°C, blue) and chlorophyll fluorescence voltages (V, green) from the CTD downcast prior to the start of the incubations.	110
Supplemental Figure 2.2: pH (total scale, left y-axis) and carbonate ion concentrations [CO ₃ ²⁻] (μmol kg ⁻¹ , right y-axis) computed from dissolved inorganic carbon (DIC) and total alkalinity (TA) for each experiment (E).	111
Supplemental Figure 2.3: Labile particulate iron concentrations (LpFe, nmol L ⁻¹) calculated from total dissolvable iron minus dissolved iron concentrations.	112
Supplemental Figure 2.4: Chlorophyll a concentrations (μg L ⁻¹) from the parallel incubations.	113
Supplemental Figure 2.5: Macronutrient concentrations (μmol L ⁻¹) from the parallel <i>in situ</i> incubations at the initial time point (T0) and after 24 hours of incubation with control (C), iron addition (+Fe), or DFB addition (+DFB) treatments.	114
Supplemental Figure 2.6: Short-term iron-59 uptake rates normalized to particulate organic carbon (μmol Fe mol C ⁻¹ hr ⁻¹) for inorganic iron (FeCl ₃ , black, left y-axis) and organically-complexed iron as ferrioxamine B (FeDFB, white, right y-axis).	115
Supplemental Figure 2.7: Short-term iron-59 uptake rates over time.	116
Supplemental Figure 2.8: Community growth rates (μ, d ⁻¹) for each treatment within each experiment.	117
Supplemental Figure 2.9: The relative abundance of diatom orders and genera within diatom annotated 18S rRNA amplicons at each experiment (E1-4).	118
Supplemental Figure 2.10: The relative abundance of bacterial classes and orders from 16S rRNA amplicons at each experiment (E1-4).	119
Supplemental Figure 2.11: PCoA plots of Bray-Curtis dissimilarity for the eukaryotic community (18S rRNA, top) and prokaryotic community (16S rRNA, bottom).	120

Supplemental Figure 2.12: The number of significantly differential abundant amplicon sequence variants (ASVs) at each experiment determined by DESeq2 between the 800 μ atm (right) and 400 μ atm (left) treatments at the final time point for each experiment ($P < 0.05$).....	121
Supplemental Figure 2.13: Venn diagrams of detected orthologous genes among the communities at the different experiments (E1-E4) for (A) centric diatoms and (B) pennate diatoms.	122
Supplemental Figure 2.14: Normalized transcript expression of iron starvation-induced proteins (ISIPs) in pennate diatoms (log ₁₀ scale): ISIP1 (red), ISIP2A (phytotransferrin or pTF, blue), ISIP3 (green).	123
Supplemental Figure 2.15: Normalized (total sum scaled) protein abundances of iron starvation-induced proteins (ISIPs) in centric diatoms (top row) and pennate diatoms (bottom row).	124
Supplemental Figure 2.16: Photosynthetic and ribosomal protein mass fractions for each Experiment (E), and taxonomic group with at least 50 unique peptides for each functional category.	125
Supplemental Figure 2.17: Maximum likelihood phylogenetic tree as inferred from iron permease (FTR1) amino acid sequences (255 taxa/241 positions) under the LG+C40+F+R8 model as implemented in IQTree 2.	126
Supplemental Figure 2.18: Maximum likelihood phylogenetic tree as inferred from CDF-family amino acid sequences (542 taxa/309 positions) under the LG+F+R10 model as implemented in IQTree 2.	127
Supplemental Figure 2.19: Maximum likelihood phylogenetic tree as inferred from FecCD-like sequences (58 taxa/250 positions) under the LG+C20+F+R5 model as implemented in IQTree 2.	128
Supplemental Figure 2.20. Maximum likelihood phylogenetic tree as inferred from SLC49 - family amino acid sequences (374 taxa/386 positions) under the LG+C40+F+R10 model as implemented in IQTree 2.	129
Supplemental Figure 2.21: Normalized transcript expression of ferrichrome binding protein 1 (FBP1) in all diatoms.	130
Supplemental Figure 2.22: Normalized (total sum scaled) protein abundances of ferrichrome binding protein 1 (FBP1) in all diatoms at Experiment 1.	131
Supplemental Figure 2.23: MA plots for centric diatoms at Experiment 1 (T72) and Experiment 3 (T48).	132
Supplemental Figure 2.24: Maximum likelihood phylogenetic tree as inferred from mitoNEET amino acid sequences (279 taxa/102 positions) under the LG+C40+F+R7 model as implemented in IQTree 2.	133
Supplemental Figure 2.25: Differential expression of nitrogen assimilation transcripts for certain taxonomic groups and experiments (E).	134
Supplemental Figure 2.26: MA plots for <i>Ostreococcus</i> at the second time point (96 hours) at Experiment 3.	135

Supplemental Figure 2.27: SLC4-family bicarbonate transporter transcript expression ($\log_{10}$ scale) in centric diatoms (top) and pennate diatoms (bottom).	136
Supplemental Figure 2.28: Rarefaction curves for eukaryotes (18S) and prokaryotes (16S) for each sample.	137
Figure 3.1: Overview of CalCOFI samples used in this study.	153
Figure 3.2: Correlations among abundance metrics for each major eukaryotic phytoplankton group.	166
Figure 3.3: Correlations among abundance metrics for cyanobacteria.	170
Figure 3.4: Productivity-diversity relationships in the region with biomass (chlorophyll <i>a</i> concentrations) as a proxy for productivity and diversity expressed as the Shannon Index (H') for both the 18S-V4 (red) and 18S-V9 (blue) data.	175
Figure 3.5: <i>Pseudo-nitzschia</i> abundances.	179
Figure 3.6: Strongly significant correlations between diagnostic pigments and gene expression (transcripts L^{-1}) associated with each respective group ($r > 0.60$, FDR < 0.05) and organized by KEGG module classes.	182
Supplemental Figure 3.1: (A) Temperature ($^{\circ}C$) and (B) nitrate concentrations ($\mu mol L^{-1}$) among seasons for the RNA samples.	200
Supplemental Figure 3.2: Cell abundances (cells mL^{-1}) for <i>Prochlorococcus</i> (Pro) and <i>Synechococcus</i> (Syn) as (A) distributions and (B) sample-by-sample comparisons where the dashed line represents a one-to-one relationship.	201
Supplemental Figure 3.3: Pearson correlations among absolute abundances of phytoplankton groups and diagnostic pigments for (A) 18S-V4 and (B) 18S-V9 data.	202
Supplemental Figure 3.4: Correlations between 18S-V4 (red) or 18S-V9 (blue) absolute abundances (copies L^{-1}) and taxon-specific contributions to total chlorophyll estimated with phytoclass ($\mu g L^{-1}$) for each major eukaryotic phytoplankton group.	203
Supplemental Figure 3.5: Linear regressions for depth (m) in predicting residuals from linear models on the absolute abundance and diagnostic pigments shown in Fig. 2B.	204
Supplemental Figure 3.6: Relative abundances based on total copies L^{-1} of Dinoflagellata ASVs on the (A) order level including Syndiniales and (B) genus level when Syndiniales are removed.	205
Supplemental Figure 3.7: Correlations between diagnostic pigment concentrations and the absolute abundances of mRNA (transcripts L^{-1}) for (A) all cyanobacteria and zeaxanthin and (B) <i>Prochlorococcus</i> and divinyl chlorophyll <i>a</i> (DVChl <i>a</i>).	206
Supplemental Figure 3.8: Productivity-diversity relationships in the region with net primary productivity (NPP) and both (a) the Shannon Index and (b) richness (ASVs) using both the 18S-V4 (red) and 18S-V9 (blue) data.	207
Supplemental Figure 3.9: Productivity-diversity relationship in the region for (A) all eukaryotic phytoplankton groups or (B) all eukaryotic groups and cyanobacteria with biomass (chlorophyll <i>a</i> concentrations) as a proxy for productivity and diversity expressed as the richness (number of ASVs) for both the 18S-V4 (red) and 18S-V9 (blue) data.	208

Supplemental Figure 3.10: Productivity-diversity relationships for each individual phytoplankton groups using richness (ASVs) against taxon-specific biomass estimated with phytoclass.	209
Supplemental Figure 3.11: Productivity-diversity relationships in the region for all eukaryotic phytoplankton groups except for dinoflagellates with biomass (chlorophyll <i>a</i> concentrations) as a proxy for productivity and diversity expressed as the Shannon Index (H') for both the 18S-V4 (red) and 18S-V9 (blue) data.	210
Supplemental Figure 3.12: Diatom-to-dinoflagellate ratio for the (A) 18S-V4 and (B) 18S-V9 data against total chlorophyll <i>a</i> concentrations.....	211
Supplemental Figure 3.13: Dinoflagellate genus-level abundances against (A) total dinoflagellate copies L^{-1} , (B) total chlorophyll <i>a</i> concentrations, and (C) peridinin concentrations.....	212
Figure 4.1: Overview of NOAA-CalCOFI Ocean Genomics (NCOG) sampling from 2014 to 2020.....	218
Figure 4.2: Overview of total diatom richness and abundances in the NCOG dataset from 2014 to 2020.....	220
Figure 4.3: Diatom abundances data against environmental measurements.	224
Figure 4.4: Genus-level diatom abundances and richness on log scales.....	227
Figure 4.5: Exponentiated Shannon diversity [$\exp(H)$] compared to environmental measurements.....	231
Figure 4.6: (A) Abundance-occupancy distributions fit with a neutral model for the V4 (left) and V9 (right) datasets.....	234
Figure 4.7: Diatom abundances and diversity in the mesopelagic zone.	237
Figure 4.8: Diatom gene expression.	241
Supplemental Figure 4.1: The abundance of diatom transcripts as (A) absolute abundances (transcripts L^{-1}) and (B) relative abundances (percent of TPM) for each depth category.	268
Supplemental Figure 4.2: Spearman correlations between different measures of diatom abundances.....	269
Supplemental Figure 4.3: 18S-V4 CLR-transformed relative abundances versus 18S-V9 CLR-transformed relative abundances for summed diatom ASVs in each NCOG sample.....	270
Supplemental Figure 4.4: Diatom cell counts from near-surface waters and pooled by regions on CalCOFI cruises from 2014 to 2020.....	271
Supplemental Figure 4.5: Summed CLR-transformed diatom abundances versus environmental variables in both V4 (red) and V9 (blue) data with Spearman rank correlation coefficients shown above each panel.	272
Supplemental Figure 4.6: Summed diatom absolute abundances (copies L^{-1}) versus environmental variables in both V4 (red) and V9 (blue) data with Spearman rank correlation coefficients.....	273
Supplemental Figure 4.7: (A) CLR-transformed relative abundances and (B) absolute abundances (copies L^{-1}) separated by season for Chlorophyta, Dinoflagellata, and Haptophyta with the V4 (red) and V9 (blue) data.....	274

Supplemental Figure 4.8: Realized temperature niches (T_{obs}) for ASVs detected in 10 or more euphotic zone samples within dominant eukaryotic phytoplankton groups for the V4 (red) and V9 (blue) data.	275
Supplemental Figure 4.9: Genus-level diatom abundances and richness (log scale) based on transcripts per L from the metatranscriptome (left) and microscopic cell counts (right, cells per 100 mL).....	276
Supplemental Figure 4.10: The percentage of rarefied reads found in near-surface versus subsurface chlorophyll maxima layers (SCML) for each diatom genus and both the 18S-V4 (left) and 18S-V9 (right) data.	277
Supplemental Figure 4.11: Relationship between genus abundance and richness as shown in Fig. 4 for the V4 (top) and V9 (bottom) data.	278
Supplemental Figure 4.12: Alpha diversity metrics between the diatom-only V4 and V9 data after rarefaction.	279
Supplemental Figure 4.13: Relationships between alpha diversity expressed as the exponentiated Shannon index and continuous variables for both the V4 (red) and V9 (blue) datasets.....	280
Supplemental Figure 4.14: Diatom diversity expressed as the (A) Shannon Index or (B) richness (number of ASVs) against taxon-specific biomass (μg chlorophyll <i>a</i> L^{-1}).....	281
Supplemental Figure 4.15: Mantel test results between distance between stations and diatom community dissimilarity within each cruise.	282
Supplemental Figure 4.16: Abundance-occupancy distributions for all (A) V4 and (B) V9 ASVs from the rarefied data.	283
Supplemental Figure 4.17: The average read abundance on a log scale per station for (A) all diatom ASVs and (B) <i>Chaetoceros</i> ASVs that are predicted to be dispersal limited by the neutral model.....	284
Supplemental Figure 4.18: Connections between (A, B) diatom V4 ASVs or (C, D) diatom V9 ASVs and other taxa as inferred from the co-occurrence network.	285
Supplemental Figure 4.19: Genus-level diatom abundances within mesopelagic zone samples from (A) rarefied relative abundances and (B) absolute 18S abundances (copies L^{-1}).	286
Supplemental Figure 4.20: The absolute abundances of diatom gene clusters.....	287
Supplemental Figure 4.21: Gene expression in other dominant phytoplankton groups.	288

LIST OF TABLES

Table 2.1: Experiment information.....	49
Supplemental Table 2.1: Reported minimum near-surface pH values in the California Current System.....	138
Supplemental Table 2.2: Carbonate system projections for Tier 1 Shared Socioeconomic Pathway (SSP) scenarios representing the wide range of uncertainty by the year 2100 (Gidden et al. 2019).	139
Supplemental Table 2.3: Average dissolved inorganic carbon (DIC) and total alkalinity (TA) as well as calculated pH, pCO ₂ , and [CO ₃ ²⁻].	140
Supplemental Table 2.4: Comparison between the number of significantly differentially expressed transcript (T) or protein (P) clusters for dominant taxonomic groups at each experiment (DESeq2 Wald Test, Benjamini & Hochberg adjusted P < 0.05).....	141
Supplemental Table 2.5: The number of significantly differentially expressed transcript (T) or protein (P) clusters for dominant taxonomic groups at each experiment (DESeq2 Wald Test, Benjamini & Hochberg adjusted P < 0.05).....	142
Supplemental Table 2.6: Liquid chromatography flow rates and solvent composition settings.	143
Supplemental Table 2.7: Mass spectrometer settings.	144
Supplemental Table 3.1: Pearson correlation coefficients for dinoflagellates with different groups removed and peridinin.	213

ACKNOWLEDGEMENTS

Thank you to my advisor, Andy Allen, for supporting my PhD, always pushing me to think bigger, and never saying no to anything. I would also like to thank my committee: Kathy Barbeau, Andrew Barton, and Karsten Zengler. Kathy, especially, made a lot of trace metal clean incubation experiments possible, only some of which are shown in Chapter 1. I am also incredibly grateful for my previous scientific mentors, Astrid Schnetzer and Adrian Marchetti, without whom, I would not be completing my PhD. Thank you also to Gilbert Bretado, who I am confident is the best graduate student affairs advisor in the world, as well as the many other incredible staff members at Scripps and JCVI.

I would also like to thank the many amazing collaborators I have had the fortune of working with during my PhD, many of whom are listed as co-authors below. Thank you to the other contributing members of the CCE LTER and CalCOFI programs, of which most of this dissertation was built upon. In particular, I thank Mark Ohman as former Lead PI of the CCE LTER program and chief scientist of the P1908 cruise where the experiments in Chapter 1 were performed. Within CalCOFI and the NCOG programs, which provide the foundation for Chapters 3 and 4, Bryce Ellman, Shonna Dovel, and Megan Roadman worked tremendously hard to make both these data and the programs at large possible. Elizabeth Venrick also contributed significantly with her extraordinary record of phytoplankton cell abundances from CalCOFI cruises.

It is no coincidence that two Allen Lab members, Hong Zheng and Ariel Rabines, are listed as co-authors on three out of four chapters. Thank you both for all your hard work, not only for these projects, but also for the lab as a whole. Thank you to Anne Schulberg, not only for the enormous number of samples collected on CalCOFI cruises, but also for your tremendous

support as a friend. Thank you also to the other members of the Allen Lab, past and present, for making up such a special team.

Thank you to the National Science Foundation Graduate Research Fellowship Program and the LaVerne Noyes Foundation for providing direct financial support of my PhD. I would also like to thank the Embassy of France for funding my stay in Paris through the Chateaubriand Fellowship Program as well as Chris Bowler and the other members of his lab for welcoming me at the ENS.

Chapter 1, in full, is a reprint of the material as it appears in Lampe, R.H., Coale, T.H., McQuaid, J.B., and Allen, A.E. in press. *Annual Review in Microbiology* 78. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in Lampe, R.H., Coale, T.H., Forsch, K.O., Jabre, L.J., Kekuewa S., Bertrand, E.M., Horák A., Oborník M., Rabines, A.J., Rowland, E., Zheng, H., Andersson, A.J., Barbeau, K.A., and Allen, A.E. 2023. *Nature Communications* 14: 7215. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, has been submitted for publication and is a reprint of the material as it appears in Lampe, R.H., Rabines, A.J., Wood, S.M., Schulberg, A., Goericke, R., Venepally, P., Zheng, H., Stukel, M.R., Landry, M.R., Barton, A.D., and Allen, A.E. 2024. *EGUsphere* [preprint] <https://doi.org/10.5194/egusphere-2024-3285>. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Lampe, R.H., Rabines, A.J., Schulberg, A., Goericke, R., Venepally, P., Sorenson, M.,

Zheng, H., Semmens, B.X., Goodwin, K.D., Barton, A.D., and Allen, A.E. The dissertation author was the primary investigator and author of this paper.

VITA

- 2012 Bachelor of Science in Computer Science, North Carolina State University
- 2018 Master of Science in Marine Science, University of North Carolina at Chapel Hill
- 2024 Doctor of Philosophy in Oceanography, University of California San Diego

PUBLICATIONS

Lampe R.H., T.H. Coale, J.B. McQuaid, and A.E. Allen. (in press) Molecular Mechanisms for Iron Uptake and Homeostasis in Marine Eukaryotic Phytoplankton. *Annual Review of Microbiology*.

James C.C., A.E. Allen, R.H. Lampe, A. Rabines, and A.D. Barton. (2024) Endemic, cosmopolitan, and generalist taxa and their habitat affinities within a coastal marine microbiome. *Scientific Reports* 14: 22408.

Haghani N., R. H. Lampe, B. Samuel, S. H. Chalasani, and M. A. Matty. (2024) Identification and characterization of a skin microbiome on *Caenorhabditis elegans* suggests environmental microbes confer cuticle protection. *Microbiology Spectrum* 12:e00169-24.

Ternon, E., M. L. Carter, L. Cancelada, R. H. Lampe, A. E. Allen, C. Anderson, K. Prather, W. Gerwick. (2023) Yessotoxin production and aerosolization during the unprecedented red tide of 2020 in southern California. *Elementa: Science of the Anthropocene*, 11:00021

Lampe, R. H., T. H. Coale, K. O. Forsch, L. J. Jabre, S. Kekuewa, E. M. Bertrand, A. Horák, M. Oborník, A. J. Rabines, E. Rowland, H. Zheng, A. J. Andersson, K. A. Barbeau, and A. E. Allen. (2023) Short-term acidification promotes diverse iron acquisition and conservation mechanisms in upwelling-associated phytoplankton. *Nature Communications* 14, 7215.

Satterthwaite, E. V., A. E. Allen, R. H. Lampe, Z. Gold, A. R. Thompson, N. Bowlin, R. Swalethorp, K. D. Goodwin, E. L. Hazen, S. J. Bograd, S. A. Matthews, and B. X. Semmens. (2023) Toward identifying the critical ecological habitat of larval fishes: An environmental DNA window into fisheries management. *Oceanography* 36 (Supplement 1), 90-93.

Maniscalco, M. A., M. A. Brzezinski, R. H. Lampe, N. R. Cohen, H. M. McNair, K. A. Ellis, M. Brown, C. P. Till, B. S. Twining, K. W. Bruland, A. Marchetti, and K. Thamtrakoln. (2022) Diminished carbon and nitrate assimilation drive changes in diatom elemental stoichiometry independent of silicification in an iron-limited assemblage. *ISME Communications* 2, 57.

- Cohen, N. R., H. A. Alexander, A. I. Krinos, S. K. Hu, and R. H. Lampe. (2022) Marine microeukaryote metatranscriptomics: sample processing and bioinformatic workflow recommendations for ecological applications. *Frontiers in Marine Science* 9, 867007.
- Coale, T. H., E. M. Bertrand, R. H. Lampe, and A. E. Allen. 2022. Molecular Mechanisms Underlying Micronutrient Utilization in Marine Diatoms, p. 567-604. *In* A. Falciatore and T. Mock [eds.], *The Molecular Life of Diatoms*. Springer International Publishing.
- James, C. C., A. D. Barton, L. Z. Allen, R. H. Lampe, A. Rabines, A. Schulberg, H. Zheng, R. Goericke, K. D. Goodwin, and A. E. Allen. (2022) Influence of nutrient supply on plankton microbiome biodiversity and distribution in a coastal upwelling region. *Nature Communications* 13, 2448.
- Kranzler, C. F., M. A. Brzezinski, N. R. Cohen, R. H. Lampe, M. Maniscalco, C. P. Till, J. Mack, J. R. Latham, K. W. Bruland, B. S. Twining, A. Marchetti, and K. Thamtrakoln. (2021) Impaired viral infection and reduced mortality of diatoms in iron-limited oceanic regimes. *Nature Geoscience* 14, 231-237.
- Lampe, R. H., G. Hernandez, Y. Y. Lin, and A. Marchetti. (2021) Representative diatom and coccolithophore species exhibit divergent responses throughout simulated upwelling cycles. *mSystems* 6:10.1128/msystems.00188-21.
- Lampe, R. H., S. Wang, N. Cassar, and A. Marchetti. (2019) Strategies among phytoplankton in response to alleviation of nutrient stress in a subtropical gyre. *The ISME Journal* 13, 2984-2997.
- Till, C. P., J. R. Solomon, N. R. Cohen, R. H. Lampe, A. Marchetti, T. H. Coale, and K. W. Bruland. (2019) The iron limitation mosaic in the California Current System: factors governing Fe availability in the shelf/near-shelf region. *Limnology and Oceanography* 64, 109-123.
- Lampe, R. H., E. L. Mann, N. R. Cohen, C. P. Till, K. Thamtrakoln, M. A. Brzezinski, K. W. Bruland, B. S. Twining, and A. Marchetti. (2018) Different iron storage strategies among bloom-forming diatoms. *Proceedings of the National Academy of Sciences USA* 115, E12275-E12284.
- Lampe, R. H., N. R. Cohen, K. A. Ellis, K. W. Bruland, M. T. Maldonado, T. D. Peterson, C. P. Till, M. A. Brzezinski, S. Bargu, K. Thamtrakoln, F. I. Kuzminov, B. S. Twining, and A. Marchetti. (2018) Divergent gene expression among phytoplankton taxa in response to upwelling. *Environmental Microbiology* 20, 3069-3082.
- Cohen, N. R., K. A. Ellis, R. H. Lampe, H. McNair, B. S. Twining, M. A. Brzezinski, M. T. Maldonado, F. I. Kuzminov, K. Thamtrakoln, C. P. Till, K. W. Bruland, W. G. Sunda, S. Bargu, and A. Marchetti. (2017) Variations in diatom transcriptional response to changes in iron availability across ocean provinces. *Frontiers in Marine Science* 4, 360.
- Cohen, N. R., K. A. Ellis, W. G. Burns, R. H. Lampe, N. Schuback, Z. I. Johnson, S. A. Sañudo-Wilhelmy, and A. Marchetti. (2017) Iron and vitamin interactions in marine diatom isolates and natural assemblages of the Northeast Pacific Ocean. *Limnology and Oceanography* 62, 2076-2096.

Marchetti, A., C. M. Moreno, N. R. Cohen, I. Oleinikov, K. DeLong, B. S. Twining, E. V. Armbrust, and R. H. Lampe. (2017) Development of a molecular-based index for assessing iron status in bloom-forming pennate diatoms. *Journal of Phycology* 53, 820-832.

Schnetzer, A., R. H. Lampe, C. R. Benitez-Nelson, A. Marchetti, C. L. Osburn, and A. O. Tatters. (2017) Marine snow formation by the toxin-producing diatom, *Pseudo-nitzschia australis*. *Harmful Algae* 61, 23-30.

ABSTRACT OF THE DISSERTATION

Environmental Genomic Investigations of Phytoplankton in the California Current Ecosystem

by

Robert H. Lampe

Doctor of Philosophy in Oceanography

University of California San Diego, 2024

Professor Andrew E. Allen, Chair

Coastal upwelling regions, such as the California Current Ecosystem, are among the most biologically productive regions in the ocean. Eukaryotic phytoplankton, particularly diatoms, underlie this high productivity, forming the base of the food web and driving biogeochemical cycling in the region. As a result, understanding how anthropogenic change, which is superimposed on the high variability of coastal upwelling regions, may influence phytoplankton abundances and physiology is critical towards forecasting the future state of the ocean. Concurrently, advancements in molecular-based approaches offer the ability to examine these organisms both within their environment and with an unprecedented level of detail.

In Chapter 1, we review, from a molecular perspective, the diverse mechanisms that marine eukaryotic phytoplankton have evolved related to the acquisition and use of the micronutrient iron. Iron frequently limits phytoplankton growth in the large areas of the ocean including the California Current, and its bioavailability for phytoplankton is expected to be reduced by ocean acidification, which is driven by rising atmospheric carbon dioxide concentrations. Therefore, in Chapter 2, we examine the effects of ocean acidification on natural phytoplankton communities in the California Current Ecosystem. We show that eukaryotic phytoplankton respond to increased acidification with a variety of iron uptake and conservation mechanisms that are indicative of increased iron stress, further supporting that acidification will negatively impact phytoplankton communities.

In Chapter 3, we describe how pigment-based approaches to measure phytoplankton abundances and community composition align with quantitative DNA- or RNA-based measurements. As phytoplankton pigments may be modeled from hyperspectral remote sensing data, these results further support that phytoplankton community composition, including harmful bloom-forming taxa, can be assessed from satellite-based ocean color observations.

In Chapter 4, these quantitative molecular approaches are leveraged to examine patterns in diatom abundances, diversity, and gene expression in the California Current Ecosystem. In contrast to other phytoplankton groups, seasonal upwelling conditions are strongly associated with increases in diatom abundances and diversity as well as differences in gene expression. Furthermore, dominant bloom-forming diatoms in the nearshore environment are shown to be dispersal limited over relatively small spatial scales, as they sink out of the euphotic zone and contribute to carbon export rather than becoming established offshore.

Collectively, these results improve our understanding of phytoplankton in upwelling regions and support our ability to predict their future responses. In some areas, upwelling may intensify, resulting in an abundance of diverse diatoms that contribute to carbon export. In other areas, increased stratification may result in a diatom community with both low abundances and low diversity. Meanwhile, increased acidification may exacerbate iron limitation leading to declines in productivity. Continued environmental genomic investigations of phytoplankton in upwelling regions that further span phases of climate modes or experiment with other projected future conditions may reveal additional patterns and responses, increasing our understanding of phytoplankton from individual genes to entire communities.

CHAPTER 1 - MOLECULAR MECHANISMS FOR IRON UPTAKE AND HOMEOSTASIS IN MARINE EUKARYOTIC PHYTOPLANKTON

Robert H. Lampe, Tyler H. Coale, Jeffrey B. McQuaid, and Andrew E. Allen

Abstract

The micronutrient iron is essential for phytoplankton growth due to its central role in a wide variety of key metabolic processes including photosynthesis and nitrate assimilation. As a result of scarce bioavailable iron in seawater, marine primary productivity is often iron-limited with future iron supplies remaining uncertain. Although evolutionary constraints resulted in high cellular iron requirements, phytoplankton evolved diverse mechanisms that enable uptake of multiple forms of iron, iron storage over short and long timescales, and modulation of their iron requirement under stress. Genomics continues to increase our understanding of iron-related proteins that are homologous to those characterized in other model organisms, while recently, molecular and cell biology is revealing unique genes and processes with connections to iron acquisition or use. Moreover, there are an increasing number of examples showing the interplay between iron uptake and extracellular processes such as boundary layer chemistry and microbial interactions.

1. Introduction

Nutrient availability in the ocean has immense influence on global primary productivity and microbially-mediated biogeochemical cycling. In this context, the role of iron (Fe) as a micronutrient that limits phytoplankton growth has garnered significant interest over the past several decades. Although reviews from the 1950s and 1960s in this very journal describe early recognition of the universal requirement for iron to support life (Provasoli 1958), uncertainty remained surrounding its significance in controlling phytoplankton growth in the natural

environment (Lewin and Guillard 1963; Strickland 1965). Following the first accurate measurements of dissolved iron in seawater in the 1980s that revealed surface concentrations less than 0.5 nmol/kg (Gordon et al. 1982; Landing and Bruland 1981; Landing and Bruland 1987) and experiments demonstrating increases in phytoplankton biomass after iron enrichment (Coale 1991; Martin and Fitzwater 1988; Martin et al. 1989), the high importance of iron as a control on phytoplankton abundances and productivity in the ocean began to be more fully recognized.

With extensive and on-going efforts to obtain high-quality iron measurements in the ocean, namely via the GEOTRACES program (Anderson 2020), dissolved iron concentrations in the surface ocean are now known to typically be 0.1 to 0.6 nM range but as low as 0.01 nM (GEOTRACES Intermediate Data Product Group 2023; Tagliabue et al. 2012) contributing to widespread limitation of growth in large areas of the ocean (Moore et al. 2001). Moreover, iron has been found to simultaneously co-limit eukaryotic phytoplankton growth with nitrogen (Browning et al. 2017), the other trace metal Mn (Browning et al. 2021), and light in subsurface chlorophyll maxima layers (Hogle et al. 2018) and in the Southern Ocean (Boyd 2002). Furthermore, there are interactive effects between iron and other micronutrients, such as vitamin B₁₂ (Bertrand et al. 2015; Bertrand et al. 2007), copper (Peers et al. 2005), and Zn (Crawford et al. 2003). The availability of silicic acid and iron may also have interactive effects on diatoms, a lineage of phytoplankton belonging to the Stramenopile clade, as they require silicic acid to produce their cell walls (Brzezinski et al. 2011; Hutchins et al. 2001; Marchetti et al. 2010).

Iron limitation in phytoplankton is traced to their early evolutionary history, most notably with the evolution of oxygenic photosynthesis. In phytoplankton, iron has essential roles as a cofactor for several proteins in diverse cellular processes, namely photosynthesis and nitrate assimilation (Marchetti and Maldonado 2016). Oxygenic photosynthesis emerged approximately

three billion years ago when dissolved iron concentrations were significantly higher, and phytoplankton made use of this abundant iron to irreversibly incorporate it into their photosynthetic electron transport chain and other key enzymes (Behrenfeld and Milligan 2013; Wade et al. 2021). Following widespread and persistent oxygenation of surface waters 1,850-1,250 Mya (Anbar and Knoll 2002), iron in seawater formed insoluble ferric oxyhydroxides that precipitated, resulting in the low dissolved iron concentrations observed today (Behrenfeld and Milligan 2013; Liu and Millero 2002). In addition to biological uptake, dissolved iron adsorbs onto particle surfaces or forms colloidal authigenic particles that may aggregate into larger particles and sink, further contributing to iron removal from the surface ocean, although these processes are somewhat reversible (Boyd et al. 2017; Tagliabue et al. 2023).

Iron removal is balanced by inputs that vary regionally (Tagliabue et al. 2017). On global scales, iron supply is dominated by continental margin sediments where lateral advection transports iron beyond coastal waters. Iron released from hydrothermal vents is also a significant source as it is transported long distances and eventually upwelled, particularly in the Southern Ocean (Ardyna et al. 2019). In contrast, atmospheric dust deposition is relatively important in low latitude regions (Tagliabue et al. 2017), and on regional scales, rivers, glaciers, and sea ice melt also contribute to iron inputs (Gerringa et al. 2012; Raiswell et al. 2008; Rijkenberg et al. 2014). Following biological uptake, rapid recycling can further sustain iron in the euphotic zone (Boyd et al. 2017).

Iron bioavailability is further influenced by its complex physicochemical speciation in seawater (Figure 1.1). It exists in two oxidation states: the more soluble ferrous iron (Fe(II)) and less soluble ferric iron (Fe(III)) that forms Fe(III) oxyhydroxide precipitates (Byrne and Kester 1976; Kuma et al. 1992). Fe(II) generally exists in relatively low concentrations as it undergoes

oxidation to Fe(III) within minutes to hours (Santana-Casiano et al. 2005), although at times, Fe(II) can account for a high percentage of total dissolved iron (Hopwood et al. 2020). Within the dissolved iron pool, less than 1% is presumed to exist as soluble inorganic Fe(II) or Fe(III), the sum of which is referred to as Fe'. The remainder is usually bound to organic ligands such as bacterial siderophores, humic acids, proteins, and exopolysaccharides. With a range of relatively strong to weak affinities for iron, these organic ligands allow dissolved iron concentrations to persist above the solubility limits of its inorganic species (Gledhill and Buck 2012).

Transformations among these chemical species are influenced by pH, temperature, oxygen, and photochemistry including the dissolution of particulate iron (Fitzsimmons and Conway 2023; Lorenzo et al. 2020; Tagliabue et al. 2009). The result is a spectrum of iron bioavailability from the most available but least abundant being Fe' to the least available generally being iron bound to certain strong organic ligands (Lis et al. 2015).

In response to iron stress or limitation, phytoplankton exhibit a range of physiological responses that in conjunction with the physiological aspects of iron uptake are reviewed by Marchetti and Maldonado (2016). Photophysiological responses to iron stress are further detailed by Behrenfeld and Milligan (2013). Briefly, intracellular iron contents, or quotas, decrease. Maximal uptake rates of both Fe' and organically-complexed iron increase suggesting the upregulation of various cell surface iron transporters (Harrison and Morel 1986; Maldonado and Price 2001). Cell sizes may also decrease thus reducing overall nutrient requirements and improving nutrient uptake kinetics (Marchetti and Maldonado 2016). Phytoplankton remodel their photosynthetic architecture, including decreases in their pigment inventories leading to observable changes in their fluorescent properties (Behrenfeld and Milligan 2013; Davey and Geider 2001; Ryan-Keogh and Thomalla 2020). In diatoms, iron limitation also affects their

elemental stoichiometry. Specifically, diatoms show increased silicon (Si) relative to carbon and nitrogen (Hutchins and Bruland 1998), which appears to be driven by reductions in cellular carbon and nitrogen rather than increases in Si (Maniscalco et al. 2022).

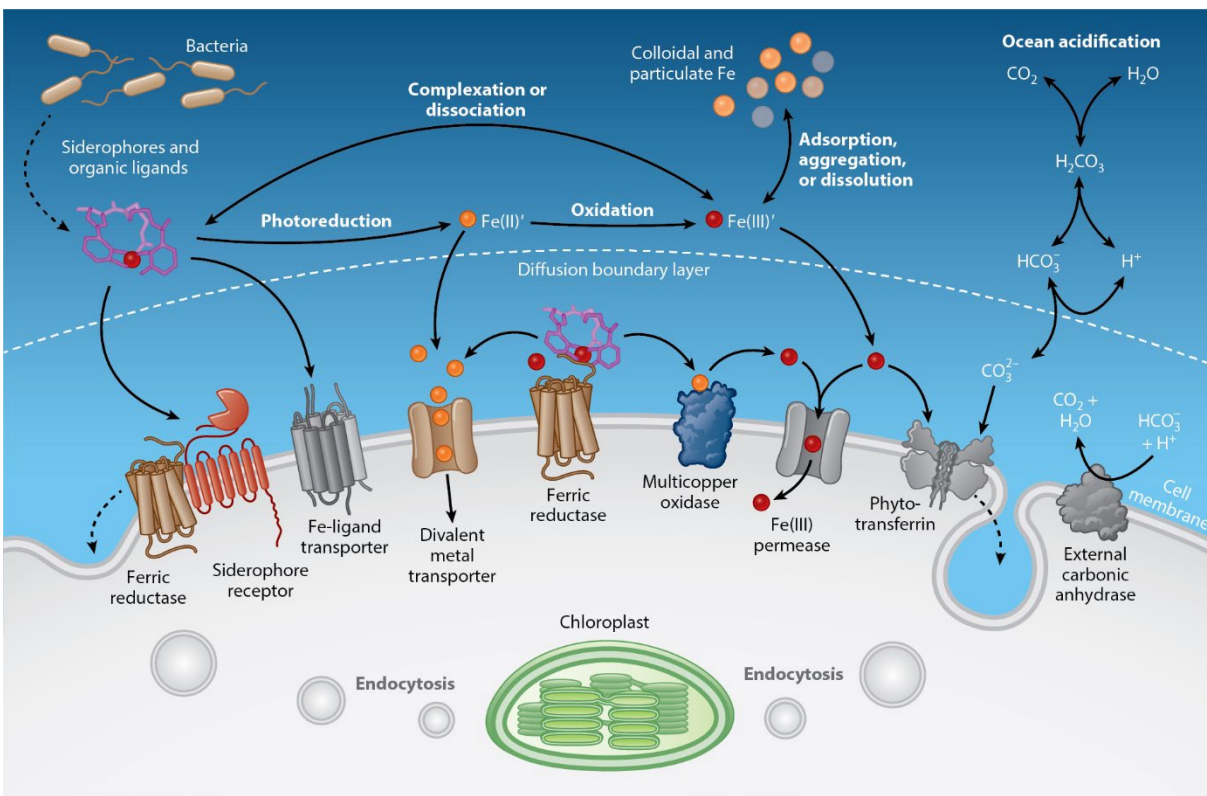


Figure 1.1: Cell surface view of iron uptake and processes that influence iron bioavailability in seawater. Figure adapted from images created with BioRender.com.

Here we describe molecular mechanisms underlying iron uptake, cellular iron demand, and responses to iron stress in marine eukaryotic phytoplankton. Due to their prominence and ecological importance, previous studies, and therefore this review, largely focus on diatoms (Bacillariophyta) that alone comprise considerable diversity (Bowler et al. 2010). However, many diatom proteins are homologous to those found in other phytoplankton lineages, suggesting that some of these mechanisms translate to different taxa. Studies on marine green

algae (Prasinophyta), haptophytes (Haptophyta), dinoflagellates (Dinoflagellata), and other eukaryotic organisms are also included.

2. Iron Uptake Mechanisms and Transporters

Corresponding to the wide variety of iron species available in seawater, eukaryotic phytoplankton evolved to possess a variety of iron uptake mechanisms enabling uptake of multiple iron sources, thus combating low iron availability and diffusion limitation (Figure 1.1)(Sunda and Huntsman 1995; Völker and Wolf-Gladrow 1999). The evolutionary history of eukaryotic phytoplankton is complex and diverse ranging from primary endosymbiotic events that created the green and red algal lineages to secondary and tertiary endosymbioses (Keeling 2010). Genetic origins within these lineages are further complicated by extensive horizontal gene transfer with bacteria (Fan et al. 2020). Diatoms, for example, are secondary endosymbionts with a mosaic of genes derived from both red and green algal endosymbionts, horizontally transferred genes of bacterial origin, and the retention of “metazoan-type” genes from their heterotrophic exosymbiont ancestor (Bowler et al. 2010). This conglomeration of genes allows certain eukaryotic phytoplankton to possess diverse functional repertoires where iron-related genes are no exception.

2.1 Inorganic Iron Uptake and Intracellular Transport

As normally less than 1% of the dissolved iron pool, soluble inorganic iron (Fe^+) exists in picomolar concentrations, yet it is the most bioavailable species and a primary source of iron for marine phytoplankton (Gledhill and Buck 2012; Morel et al. 2008). Inorganic iron uptake systems characterized in terrestrial organisms have affinities for iron in the micromolar range

suggesting that marine phytoplankton must have novel mechanisms for inorganic iron acquisition under low concentrations (Sutak et al. 2020). One such mechanism is the protein phytoferritin (pTF), an outer membrane receptor that allows cells to import individual Fe(III) atoms via endocytosis (McQuaid et al. 2018)(Figure 1.1). This protein family includes the genes previously known as *ISIP2A* in diatoms (Allen et al. 2008) and *FEA1* in green algae (Lelandais et al. 2016) that are found widely across phytoplankton lineages (Morrissey et al. 2015) and upregulated under low iron availability (Bender et al. 2018; Cohen et al. 2017b).

The evolution of pTF is a remarkable example of convergent evolution with transferrins that are found in metazoans and certain photosynthetic organisms, such as land plants and the green algae *Dunaliella salina*. Specifically, both families convergently evolved a requirement for carbonate (CO_3^{2-}) to coordinate Fe(III) ion binding resulting in a second-order dependence on the concentrations of both Fe' and CO_3^{2-} (Baker et al. 2003; McQuaid et al. 2018). These second-order kinetics allow pTF to leverage high concentrations of CO_3^{2-} to acquire trace quantities of iron; however, in environments with low iron concentrations, carbonate ion concentrations may also be insufficient leading to iron-carbonate co-limitation with respect to pTF-mediated iron uptake (McQuaid et al. 2018).

Additionally, carbonate ion concentrations are declining due to ocean acidification (Gruber et al. 2019), further decreasing the ability to bind inorganic iron. Acidification also makes iron less likely to dissociate from certain organic ligands, further lowering concentrations of Fe' (Shi et al. 2010). Collectively, these effects may explain the negative impacts of acidification on iron-stressed phytoplankton (Lampe et al. 2023; Sugie et al. 2013; Trimborn et al. 2017), although changes in both pH and carbonate chemistry at the cell surface via

photosynthesis and extracellular enzymes, such as carbonic anhydrases may allow phytoplankton to overcome some of these effects (Liu et al. 2022) (Figure 1.1).

Prior to the discovery of pTF, kinetic models for iron uptake suggested that dissolved iron species, including soluble inorganic ferric iron, or $\text{Fe(III)}'$, required reduction at the cell surface (Shaked et al. 2005). This model was based on the reductive iron uptake system in baker's yeast (*Saccharomyces cerevisiae*) and supported by similar genes found in the genome of the diatom *Thalassiosira pseudonana* (Armbrust et al. 2004), as well as physiological experiments with *Thalassiosira spp.* (Maldonado et al. 2006). This system includes cell surface ferric reductases that generate free ferrous iron, or $\text{Fe(II)}'$, from both $\text{Fe(III)}'$ and organically complexed iron. $\text{Fe(II)}'$ is then re-oxidized by a multi-copper oxidase that is coupled to a ferric iron permease (FTR) for transport across the cell membrane (Figure 1.1). This process enables selective iron transport, thus avoiding the uptake of unwanted metals (Blaby-Haas and Merchant 2012); however, as previously described, the affinities for iron uptake with this system are likely in the micromolar range (Lesuisse and Labbe 1989). While experiments suggest that diatoms exhibit extracellular Fe(II) oxidation that is impacted by copper availability (Maldonado et al. 2006), it is unknown whether or not this system is viable and functions at the cell surface as opposed to within endosomes. Additionally, while FTRs in *T. pseudonana* are localized to the cell surface, the localization of the multicopper oxidase remains unclear, leading to the alternative hypothesis that FTR functions as an independent transporter for $\text{Fe(III)}'$ uptake (A.B. Kustka, personal communication).

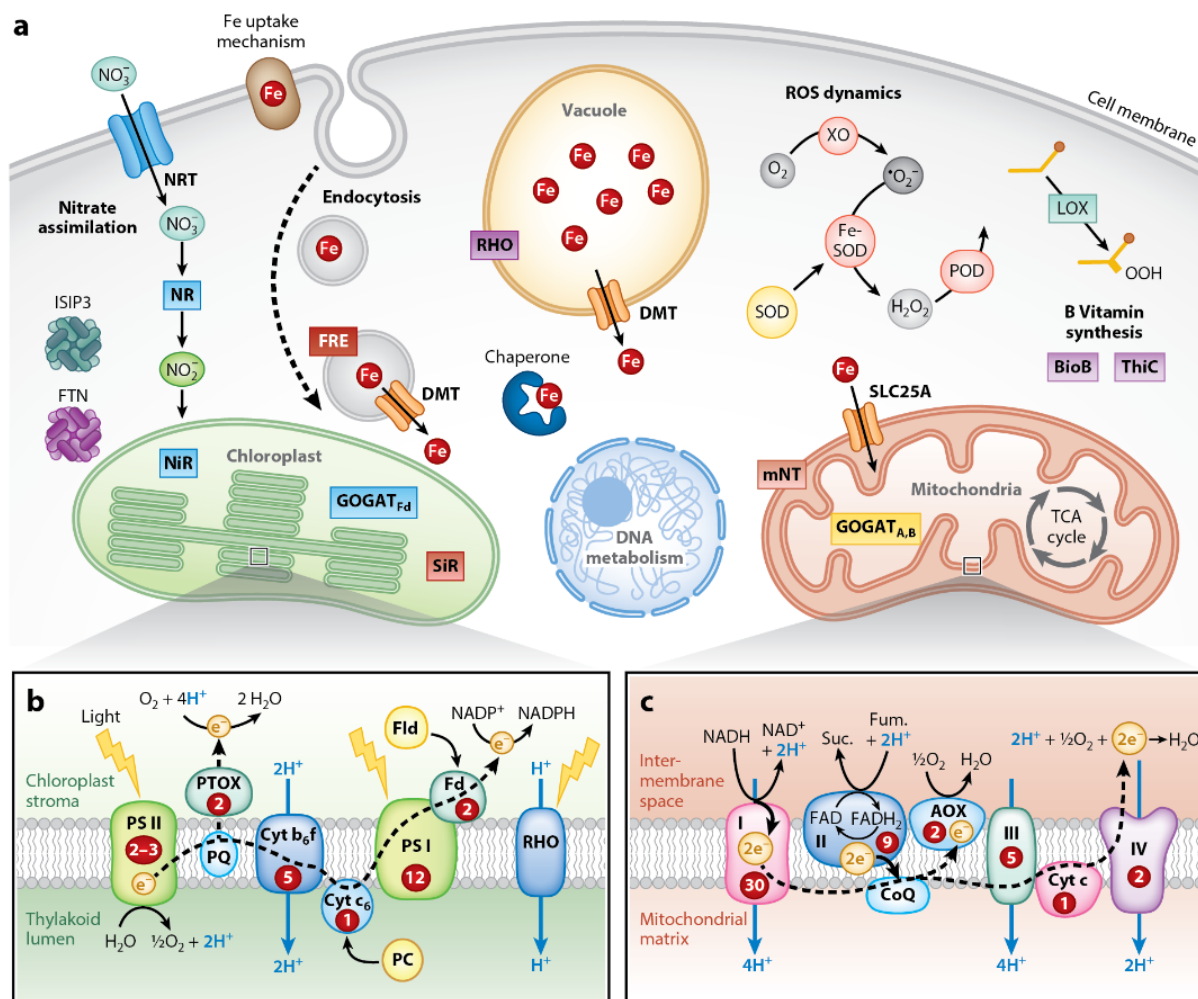


Figure 1.2: (a) Intracellular view of iron transport and storage as well as enzymes associated with iron demand. (b) Iron requirements for photosynthetic electron transport. (c) Iron requirements for mitochondrial electron transport. Abbreviations: AOX, alternative oxidase; BioB, biotin synthase; cyt, cytochrome; DMT, divalent metal transporter; FAD, flavin adenine dinucleotide; Fd, ferredoxin; Fld, flavodoxin; FRE, ferric reductase; FTN, ferritin; GOGAT, glutamate synthase; ISIP3, iron starvation-induced protein 3; LOX, lipoxygenase; mNT, mitoNEET; NR, nitrate reductase; NiR, nitrite reductase; NRT, nitrate transporter; PC, plastocyanin; POD, peroxidase; PQ, plastoquinone pool; PS, photosystem; PTOX, plastid terminal oxidase; RHO, xanthorhodopsin; ROS, reactive oxygen species; Suc, succinate; SiR, sulfite reductase; SLC25A, solute carrier family 25; SOD, superoxide dismutase; TCA, tricarboxylic acid; ThiC, phosphomethylpyrimidine synthase; XO, xanthine oxidase. Figure adapted from images created with BioRender.com

Direct uptake of inorganic ferrous iron, or Fe(II)', has been observed in different microalgae (Sutak et al. 2012) and likely occurs via non-specific divalent metal transporters

(Marchetti and Maldonado 2016) that may also take up iron following ferric reductase activity (Sutak et al. 2020)(Figure 1.1). Several families of proteins may perform this function including ZIP (Zrt, Irt-like Protein), NRAMP (Natural Resistance-Associated Macrophage Protein), and CDF (Cation Diffusion Facilitator) family transporters where they have been characterized in non-marine model organisms (Blaby-Haas and Merchant 2012). Once intracellular, iron must be trafficked to different sites and transported across organelle membranes; therefore, transporters in these families may also perform this function (Figure 1.2). To transport iron across the mitochondrial inner membrane, phytoplankton likely use proteins in the mitochondrial carrier family (SLC25A28 and SLC25A37, also referred to as mitoferrins) (Blaby-Haas and Merchant 2012). Phytoplankton also possess homologs in the Ccc1/VIT1 (Ca(II)-sensitive Cross-Complementer 1 / Vacuolar Iron Transporter 1) family that transport intracellular iron in fungi and plants; however, while *VIT1* transcripts are iron-responsive in green alga *Ostreococcus tauri* (Lelandais et al. 2016), they do not appear to be so in diatoms (Lampe et al. 2018). Rather, VIT1 in diatoms has been proposed to serve a role in cadmium detoxification (Brembu et al. 2011), although its specific function remains unclear. Lastly, diatoms and the pelagophyte, *Pelagomonas calceolata*, possess ferroportins that export iron and other divalent metals in multicellular organisms (Mitchell et al. 2014; Ward and Kaplan 2012); in microalgae, they are hypothesized to export iron from endosomes (Guérin et al. 2022).

2.2 Organic Iron Uptake

Diverse eukaryotic phytoplankton isolates and natural populations also have the ability to take up iron complexed to organic ligands (Hutchins et al. 1999; Lis et al. 2015; Maldonado et al. 2005; Shaked et al. 2020). These include a variety of structurally different compounds such as polysaccharides, proteins, porphyrins (heme), humic acids, and siderophores. Collectively, these

ligands increase the solubility of iron and form iron-ligand complexes that normally represent a majority of the dissolved iron pool (Gledhill and Buck 2012).

Dozens of structurally distinct siderophores have been observed and quantified in seawater; however, the chemical structures of many ligands remains unknown (Moffett and Boiteau 2024). Rather, organic speciation of dissolved iron in seawater is usually characterized in bulk terms by concentrations and conditional stability constants that are assigned to ligand classes: L_1 for strong ligands or L_2 for weaker ligands (Gledhill and Buck 2012). Stability constants, however, are poor indicators of bioavailability, and the disassociation of iron from organic ligands to Fe' is strongly influenced by light, further complicating their bioavailability (Shaked et al. 2020) (Figure 1.1). Furthermore, microbial production of iron-binding ligands in seawater is highly heterogenous (Moffett and Boiteau 2024). As a result, it is unlikely that there is a single uptake strategy that phytoplankton can rely upon in all situations, and selective pressure exists for multiple uptake mechanisms for organically complexed iron in addition to those for inorganic iron uptake.

The mechanisms for uptake of iron complexed with polysaccharides, proteins, and humic acids remain unknown. In the case of heme uptake, SLC49-family transporters have been proposed to serve this role as in metazoans (Lampe et al. 2023). Alternatively, iron may be released at the cell surface by decycling heme oxygenases (Allen et al. 2008); however, these processes remain to be experimentally validated.

In contrast, uptake of siderophore-bound iron is more well-characterized and appears to be common among eukaryotic phytoplankton even though no marine eukaryote is known to produce them. As siderophore production is primarily attributed to bacteria for their own uptake mechanisms (Manck et al. 2022), siderophore uptake by eukaryotic phytoplankton represents

another example of phytoplankton–bacterial interactions. The diatom *Phaeodactylum tricornutum* can take up the hydroxamate siderophores ferrioxamine B (FOB) and ferrioxamine E (FOE)(Soria-Dengg and Horstmann 1995) while *Thalassiosira oceanica* can only use FOB (Maldonado and Price 2001). Many other diatoms have also demonstrated to FOB use (Strzepek et al. 2011), as well as the dinoflagellate *Amphidinium carterae* (Malych et al. 2022) and the haptophytes *Gephyrocapsa huxleyi* (formerly *Emiliana huxleyi*) and *Phaeocystis antarctica* (Mausz et al. 2020; Segovia et al. 2017; Strzepek et al. 2011).

As previously described, early studies hypothesized that eukaryotic phytoplankton relied on a yeast-like reductive iron uptake system for both inorganic and organically complexed iron (see Section 2.1)(Shaked et al. 2005). Under this model, membrane-bound ferric reductase proteins liberate iron from extracellular siderophores and other ligands to increase the Fe²⁺ concentration available for uptake at the cell surface (Maldonado et al. 2006)(Figure 1.1). In addition to possessing genes that are similar to those of the yeast-like system (Armbrust et al. 2004), many phytoplankton species are known to exhibit extracellular ferric reductase activity that increases alongside uptake rates at low iron concentrations (Shaked et al. 2005). Furthermore, Fe(II)-specific ligands, e.g. ferrozine, inhibit iron uptake, presumably by trapping reduced iron as it is produced at the cell surface. However, as this system primarily operates with micromolar-level iron concentrations (Lesuisse and Labbe 1989), it remains unclear whether it is viable for organically complexed iron at the sub-nanomolar concentrations that are typical of the marine environment. Moreover, some marine phytoplankton lack extracellular ferric reductases, including the prasinophytes *Micromonas pusilla* and *O. tauri*, the alveolate *Chromera velia*, and *G. huxleyi*, which is still capable of FOB uptake (Sutak et al. 2012; Sutak et al. 2010).

On the other hand, non-reductive siderophore uptake employing siderophore receptors occurs at lower concentrations (Lesuisse et al. 2001). Rather than employing extracellular reduction, studies with *P. tricornutum* revealed that uptake of hydroxamate siderophores, including FOB and ferrichrome, relies on siderophore binding at the cell surface followed by endocytosis (Figure 1.1). This uptake strategy was shown to involve the proteins ISIP1 (Kazamia et al. 2018), FBP1, and FRE2 (Coale et al. 2019). The precise role of ISIP1 remains unclear, but it appears to predominantly be a diatom-specific protein (Kazamia et al. 2018). FBP1 is a siderophore receptor that was acquired via horizontal gene transfer from bacteria, while FRE2 is a ferric reductase of eukaryotic ancestry (Coale et al. 2019). Orthologs of *FBP1* are found in species known to utilize FOB, including diatoms, haptophytes, dinoflagellates, and prasinophytes.

This uptake system allows phytoplankton to access iron bound to certain hydroxamate siderophores, but not catecholates or iron bound to weaker ligands such as humic substances (Coale et al. 2019). Moreover, the *FBP1* gene has diversified and duplicated within genomes; some diatom species encode up to five paralogs that potentially bind different substrates. The FRE2 protein is one of five encoded in the *P. tricornutum* genome, yet none of the others provide functional redundancy. FRE2 also does not contribute significantly to extracellular iron reduction, indicating that this function is performed by other ferric reductases that remain to be characterized. Thus, siderophore assimilation by marine eukaryotes is likely achieved through specific receptor-mediated uptake, followed by endocytosis, reduction, and dissociation the Fe-ligand complex within the cell (Figures 1 and 2). As pTF-mediated iron uptake also depends on endocytosis, these independent pathways likely converge intracellularly to deliver iron to the chloroplast where cellular iron demand is likely highest (Turnšek et al. 2021) (Figures 1 and 2).

3. Iron-containing Proteins and Storage Mechanisms

3.1. Iron as a Cofactor

The unique electrochemical properties of iron make it an ideal enzyme cofactor for electron transfer and catalysis (Balk and Schaedler 2014). As such, it primarily exists as iron-sulfur (Fe-S) clusters, heme, and di-iron or mononuclear iron in a variety of proteins relating to diverse metabolic processes (Figure 1.2).

Most notably, iron is heavily used in the photosynthetic electron transport chain (Figure 1.2b). Within microalgae, theoretical estimations and biochemical data in iron-limited diatoms suggest that photosynthesis may account for greater than 50% of cellular iron (Raven 1990; Strzepek and Harrison 2004). In linear (non-cyclic) electron flow, electrons are transferred from PSII (2-3 Fe) to the cytochrome *b₆f* complex (Cyt *b₆f*, 5 Fe) via plastoquinol (PQH₂), to PSI (12 Fe) via cytochrome *c₆* (also known as cytochrome *b₅₅₃*, 1 Fe), and lastly to a mobile ferredoxin (Fd, 2 Fe), totaling 22-23 iron atoms per linear electron transport chain (Behrenfeld and Milligan 2013; Raven et al. 1999). Phytoplankton also possess the heme-requiring protein cytochrome *c₅₅₀* encoded by the *psbV* gene, which acts as an extrinsic subunit of PSII (Roncel et al. 2012).

Similarly, iron is used in the mitochondrial electron transport chain (Figure 1.2c). There are 47 iron atoms in each: Complex I (NADH:ubiquinone oxidoreductase) has 30 Fe, Complex II (succinate dehydrogenase) has 9 Fe, Complex III (cytochrome *bc₁*) has 5 Fe, cytochrome *c* has 1 Fe, and Complex IV (cytochrome *c* oxidase) has 2 Fe, further contributing to cellular iron demand, although to a lesser extent than photosynthesis (Raven 1990; Raven and Beardall 2017). The mitochondrial alternative oxidase (AOX) is also localized to the inner mitochondrial membrane and interrupts electron transport to Complex III to mitigate stress (Murik et al. 2019); it contains 2 Fe atoms (Berthold and Stenmark 2003). Other iron-requiring mitochondrial

proteins include the tricarboxylic acid (TCA) cycle protein aconitase that isomerizes citrate to isocitrate (Read et al. 2021) and mitoNEET, which, in metazoans, resides in the mitochondrial outer membrane to regulate mitochondrial iron homeostasis (Lampe et al. 2023; Lee et al. 2022) (Figure 1.2).

In addition to electron transport chains, nitrogen assimilation requires iron and likely contributes to a large proportion of cellular iron demand (Raven 1988) (Figure 1.2). In particular, the reduction of nitrate depends on the iron-containing enzymes nitrate reductase and nitrite reductase. Furthermore, nitrite reductase and the plastid-localized glutamate synthase depend on ferredoxin for electron transfer (Pierella Karlusich et al. 2014; Smith et al. 2019). Phytoplankton also possess a mitochondrial NADPH-dependent glutamate synthase, and diatoms possess a NADPH-dependent nitrite reductase in addition to the ferredoxin-dependent gene; however, both of these enzymes also require iron (Brown et al. 2009; Smith et al. 2019; Vanoni and Curti 1999). Besides photosynthetic electron transport and nitrogen assimilation, ferredoxins are also employed in diverse other processes, including sulfite reduction via sulfite reductase (Pierella Karlusich et al. 2014).

Iron is also required for the synthesis of certain organic compounds (Figure 1.2). Among the final steps of heme synthesis are coproporphyrinogen oxidase that possesses an iron-sulfur cluster and ferrochelatase that catalyzes the insertion of iron into protoporphyrin IX yielding heme (Ferreira et al. 1995; Layer et al. 2002; Oborník and Green 2005). Cytochrome P450s are a diverse family of heme-containing enzymes that are found across the tree of life and catalyze a wide range of reactions (Gilardi and Di Nardo 2017; Teng et al. 2019); of particular note are specific carotenoid hydroxylases in the cytochrome P450 family that are involved in the synthesis of certain accessory pigments, such as fucoxanthin and other photoprotective

xanthophylls in phytoplankton (Cui et al. 2019). The synthesis of both vitamin B₁ (thiamine) and vitamin B₇ (biotin) also requires iron via the proteins phosphomethylpyrimidine synthase (ThiC) and biotin synthase (BioB), respectively (Coale et al. 2022; Cohen et al. 2017b; Croft et al. 2006; Martinez-Gomez and Downs 2008). Lastly, the synthesis of the chromophore retinal for microbial rhodopsins (see Section 4) includes a β -carotene cleavage enzyme (β -carotene 15,15'-dioxygenase) that is likely an iron-containing enzyme (Lindqvist and Andersson 2002; Strauss et al. 2023).

Some of the proteins involved in both the production and removal of reactive oxygen species (ROS) require iron as well. Xanthine oxidase is one of several oxidoreductases that produce ROS in peroxisomes (Diaz and Plummer 2018; Nishino and Okamoto 2000). Detoxification of ROS is catalyzed by superoxide dismutases (SOD) which produce hydrogen peroxide (Fridovich 1998), and one type of SOD among others that phytoplankton possess requires iron (Groussman et al. 2015). Catalases or heme peroxidases, which also require iron, then dispose of the hydrogen peroxide produced SODs (Geider and La Roche 1994).

Other iron-requiring proteins include a variety of DNA metabolism enzymes (Puig et al. 2017) and lipoxygenases (LOXs) that are responsible for the oxidation of polyunsaturated fatty acids (D'Ippolito et al. 2009)(Figure 1.2). Furthermore, there are likely additional iron-requiring metalloenzymes and processes that remain to be uncovered. In diatoms for example, iron is incorporated into their silica cell walls, yet this process is not well understood (Ellwood and Hunter 2000; Ingall et al. 2013).

3.2 Iron Detoxification and Storage

Free iron is toxic as it produces ROS via the Fenton reaction (Halliwell and Gutteridge 1992); therefore, organisms require mechanisms to safely mobilize or sequester it within a cell. For intracellular trafficking and short-term storage, phytoplankton may use phytochelatins and metallothioneins as metallochaperones, both of which are low-molecular-weight proteins rich in cysteine enabling high metal-binding capacities (Blaby-Haas and Merchant 2023; Lelandais et al. 2016; Smith et al. 2016). In the chlorarachniophyte *Bigeloviella natans* an iron-responsive CobW domain-containing protein may serve this role (Kotabova et al. 2021).

The ability to take up more iron than needed to sustain maximum growth, or luxury uptake, can provide a competitive advantage to phytoplankton. In iron-limiting regions of the ocean where iron inputs are episodic, luxury uptake coupled to long-term storage may support continued growth and allow phytoplankton to subsist, particularly for certain diatoms (Lampe et al. 2018; Twining and Baines 2022). Two long-term iron storage mechanisms in phytoplankton have been described: vacuoles and the protein ferritin (Figure 1.2).

Highly localized intracellular concentrations of iron in the diatom *Thalassiosira weissflogii* and green alga *D. salina* are indicative of a vacuolar storage mechanism (Nuester et al. 2012; Paz et al. 2007). In *D. salina*, these were further identified to be acidic vacuoles and may be analogous to those in other organisms where polyphosphate provides high iron-binding capacities. This hypothesis is further supported by the presence of the acidic vacuole-associated protein H⁺-pyrophosphatase in *D. salina* (Paz et al. 2007), which was also found to be upregulated under iron stress in *O. tauri* (Lelandais et al. 2016).

Divalent metal transporters localized to the vacuole may serve to export iron to the cytoplasm. Specially, ZIP-family and NRAMP proteins are hypothesized to serve this role in

diatoms (Lampe et al. 2018), and in *T. pseudonana*, one ZIP-family protein showed vacuolar localization (Shire 2022) (Figure 1.2). In *Ostreococcus*, high expression of *Ccc1/VIT1*, which is homologous to a vacuolar iron transporter in yeast and plants, suggests that this gene may serve this role in green algae (Lelandais et al. 2016).

Ferritins are ubiquitous proteins that assemble into nanocages capable of storing relatively high amounts of Fe(III) atoms (Marchetti et al. 2009). Diatom ferritins are generally distinct from other eukaryotic ferritins and more similar to cyanobacterial ferritins where a role in long-term storage has been shown (Groussman et al. 2015; Shire and Kustka 2015); however, this role is not ubiquitous. Coupled to their unusually high iron storage capacity, the diatom genus *Pseudo-nitzschia* thus far appears to uniquely express ferritin in response to high iron availability, and some diatoms, such as *T. pseudonana*, lack the gene completely (Cohen et al. 2018b; Lampe et al. 2018). In other diatoms, ferritin likely supports iron recycling over diel cycles, as shown in the green alga *Ostreococcus* (Botebol et al. 2015); however, diel expression of ferritin has not yet been observed in diatoms (Smith et al. 2016).

The protein ISIP3 is another potentially related storage mechanism as it has a conserved domain belonging to the ferritin superfamily (DUF305) (Behnke and LaRoche 2020). It has been found across a range of phytoplankton lineages (Behnke and LaRoche 2020) and is strongly upregulated under iron stress in diatoms (Allen et al. 2008; Chappell et al. 2015), *P. calceolata* (Guérin et al. 2022), and the haptophytes, *G. huxleyi* and *P. anartica* (Bender et al. 2018; Lampe et al. 2021), indicating widespread use among phytoplankton taxa. However, as high expression in low-iron conditions is inconsistent with a typical storage function, ISIP3 is hypothesized to have a role in iron reallocation and trafficking under low-iron conditions.

4. Cellular Adaptations to Cope with Iron Stress

In response to iron stress, phytoplankton simultaneously reduce their cellular iron requirements and improve their nutrient uptake kinetics with strategies such as downregulating iron-containing proteins and upregulating iron transporters (Marchetti and Maldonado 2016). Thus, the cellular costs associated with iron-containing proteins and membrane space available for iron uptake form the basis for iron limitation in phytoplankton (McCain et al. 2021). Upregulation of iron transporters and downregulation of iron-containing proteins such as those involved in photosynthesis and nitrate assimilation can be observed in the transcriptomes and proteomes of phytoplankton in both laboratory (Allen et al. 2008; Cohen et al. 2018a; Lampe et al. 2021; Li et al. 2020; Nunn et al. 2013) and field environments (Cohen et al. 2017b; Jabre et al. 2021; Marchetti et al. 2012; Wu et al. 2019). Reducing cellular iron requirements can be a temporary response to low iron availability or permanent adaptation for phytoplankton existing within chronically low iron areas, as demonstrated by chronically iron-limited oceanic phytoplankton having lower iron quotas and higher iron use efficiencies compared with coastal phytoplankton (Marchetti and Maldonado 2016; Marchetti et al. 2012).

In conjunction with the high iron requirement of photosynthetic proteins, a prominent manifestation of iron stress is remodeling of the photosynthetic apparatus. As the iron requirements of PSI and Cyt b₆f are higher than that of PSII (Figure 1.2b), iron-limited oceanic phytoplankton have generally reduced their PSI and Cyt b₆f, resulting in higher PSII:PSI ratios (Strzepek and Harrison 2004), but this ratio may differ among species (Strzepek et al. 2019). Furthermore, coastal species may also preferentially downregulate PSI and Cyt b₆f under iron-limited conditions (Strzepek and Harrison 2004). A reduction in PSI relative to PSII may increase a cell's reliance on the plastoquinol oxidase (PTOX) that diverts electrons following

PSII (Behrenfeld and Milligan 2013) (Figure 1.2b). This pathway may allow ATP production to continue as a proton gradient is created by PSII proton release during water splitting, with lower iron costs than linear electron transport as PTOX only requires 2 iron atoms. Overall, phytoplankton become chlorotic under iron stress—that is, pigment inventories and chlorophyll decrease (Behrenfeld and Milligan 2013; Geider et al. 1993). However, phytoplankton also increase photoprotective xanthophyll pigments and associated light-harvesting complex (LHC) proteins, such as those in the Lhcx and Lhcz families (Geider et al. 1993; Kolody et al. 2022; Smith et al. 2016). As a consequence of this remodeling, iron-limited phytoplankton have reduced PSII photochemical efficiencies, commonly measured as F_v/F_m (Behrenfeld and Milligan 2013).

Another strategy to reduce iron requirements is by using functionally equivalent proteins that do not rely on iron as cofactors. In the photosynthetic electron transport chain, plastocyanin (PC) substitutes for cytochrome c_6 as the electron transfer agent between the cytochrome b_6f complex and PSII (Figure 1.2b) (Peers and Price 2006) and may be either the sole protein serving this functional role (Scheiber et al. 2019), constitutively expressed (Marchetti et al. 2012), or upregulated in response to iron availability (Cohen et al. 2017a). As previously described, ferredoxin is used in the photosynthetic electron transport chain and nitrogen assimilation. Under iron-limitation, ferredoxin can be substituted with flavodoxin that is less-efficient but uses flavin mononucleotide rather than iron as a cofactor (La Roche et al. 1993) (Figure 1.2b). While the ratio of expression between the ferredoxin and flavodoxin correlates well with iron in many phytoplankton, it does not in diatoms (Carradec et al. 2018). This difference is explained by the possession of two distinct flavodoxins by diatoms and certain dinoflagellates with plastids of diatom origin, or dinotoms (Graff van Creveld et al. 2023). Clade

II flavodoxins are iron-responsive and substitute for ferredoxin, but Clade I flavodoxins are not iron-responsive and appear to mitigate oxidative stress. As another substitution, phytoplankton may exchange the iron-using SOD for other types of SODs that may be functionally equivalent and use different metal cofactors: Mn, Cu-Zn, and Ni (Groussman et al. 2015) (Figure 1.2).

To circumvent chlorophyll-based photosynthesis altogether, iron-limited phytoplankton may use an alternative light-harvesting protein, xanthorhodopsin (Strauss et al. 2023). As a microbial (type 1) rhodopsin, xanthorhodopsins absorb light via the chromophore retinal and may transport protons in the plastid. This proposed function would create a proton gradient driving chloroplast-localized ATP synthase and increase cellular fitness under low-iron conditions (Figure 1.2b). Alternatively, xanthorhodopsin was localized to vacuolar membranes in the polar diatom *Pseudo-nitzschia subcurvata* suggesting that it pumps protons to support vacuolar-type ATPases (Andrew et al. 2023; Sunda and Marchetti 2024).

With altered photosynthetic activity under iron stress, light-independent Calvin–Benson cycle genes at times are upregulated (Allen et al. 2008; Behrenfeld and Milligan 2013). Perhaps most consistently, phytoplankton upregulate class I fructose 1,6-bisphosphate aldolase (FBA)(Coale et al. 2022). Phytoplankton possess multiple FBAs that likely serve different functions, and in the case of the class I FBA, it may be useful in producing glyceraldehyde-3-phosphate from fructose biphosphate to temporarily fuel glycolysis (Coale et al. 2022). Respiration is also impacted by iron stress, as shown by decreased respiration rates and downregulation of associated transcripts (Allen et al. 2008; Davey and Geider 2001).

5. Iron-related Signaling and Regulation

Iron deficiency stimulates changes in the abundance of a relatively large number of transcripts and proteins, including some of the aforementioned genes related to uptake, storage, and iron-sparing measures. Additionally, iron stress has been shown to trigger programmed cell death linked to the activity of certain metacaspases (Bidle and Bender 2008). Therefore, eukaryotic phytoplankton likely possess a sensing mechanism and associated signaling pathways in relation to their iron status (Hanikenne et al. 2009). As in photosynthetic organisms at large, this putative sensing mechanism remains unknown, and only a relatively small number of potential regulatory elements have been proposed.

Cis-regulatory elements related to iron have been identified via genomic analyses of the diatoms *P. tricornutum*, *T. oceanica*, and *Fragilariopsis cylindrus*. Among these diatoms, there is a conserved palindromic motif upstream of the iron-responsive genes *ISIP1*, *FLDA*, and a class I FBA (*FBA3*) (Lommer et al. 2012) that is homologous to an iron-responsive cis-element in the model freshwater green alga *Chlamydomonas reinhardtii* (Deng and Eriksson 2007; Yoshinaga et al. 2014). In conjunction with another motif in *P. tricornutum*, genetic substitution disrupted activity of *FLDA* and *FBP1*, while *ISIP1* was disrupted with solely by one (Yoshinaga et al. 2014). Perhaps uncoincidentally, three pairs of iron-responsive genes in *P. tricornutum* have been found to be co-located in the genome: *FRE2* and *FBP1*, *CREG1* and *ISIP2B*, and *TYR1* and *ISIP1* (Allen et al. 2008). As previously described, *FRE2* and *FBP1* functionally interact to perform hydroxamate siderophore uptake (Coale et al. 2019). The timing of expression of these genes was also similar under acclimation to low-iron conditions further suggesting that expression of these gene clusters is controlled by shared *cis*-elements (Yoshinaga et al. 2014).

CREG (cellular repressor of E1A-stimulated genes) proteins have further been found to be iron-responsive in diatoms (Behnke et al. 2023; Lampe et al. 2023; Lommer et al. 2012) and the chlorarachniophyte *B. natans* (Kotabova et al. 2021). The related human CREG1 protein is a secreted glycoprotein that mediates endocytosis of a specific growth factor (Han et al. 2009). As diatom CREG proteins have been found to colocalize with pTF, CREG is hypothesized to positively regulate pTF-mediated endocytosis during iron-limited conditions (Turnšek et al. 2021).

The increased expression of certain transcription factors (TFs) under low iron suggests that they may regulate other iron-responsive genes. In *P. tricornutum*, 15 TFs of diverse families were found to be induced by low iron (Smith et al. 2016). Two of the fifteen are sigma factors that likely are localized to the chloroplast and regulate genes in the chloroplast genome. In contrast, *T. pseudonana* and *T. oceanica* had fewer differentially expressed TFs under iron limitation (Behnke et al. 2023; Goldman et al. 2019).

Post-translational regulation via protein phosphorylation, which may activate or deactivate proteins during iron stress has also been studied in the diatom *P. tricornutum* (Tan et al. 2020). Approximately 100 proteins showed phosphorylation associated with low iron conditions; 12 of these proteins were also found to transcriptionally increase under low iron. Of note, clade II flavodoxins were phosphorylated under low iron conditions suggesting phosphorylation-based activation of these proteins. pTF also displayed different phosphorylation patterns depending on time of day and iron status suggesting that phosphorylation has a regulatory role for pTF activity.

Future Issues

1. Laboratory and environmental omic studies continue to highlight the diverse and unique adaptations used by phytoplankton in response to iron availability; however, functional characterization of genes is still in its infancy, with a large number of proteins that are not well studied or are of unknown function, necessitating continued development or use of forward and reverse genetic systems with model phytoplankton species.
2. Eukaryotic phytoplankton certainly access diverse iron species, including iron bound to both known and unknown organic ligands, yet gaps in knowledge remain with respect to the molecular bases for uptake of both inorganic and organically complexed iron, including endocytosis.
3. The identification of endocytosis as a prominent mechanism for iron acquisition demands a re-evaluation of localizations of iron transporters and reductase activity. Furthermore, there are likely trade-offs to consider between high-affinity receptor-mediated endocytosis and low-affinity transporters, such as increased susceptibility to viral infection via endocytosis.
4. Not only are future iron supplies uncertain, but the interactive effects of iron with other potential stressors that can influence iron bioavailability and demand, such as ocean acidification and warming require further examination. Furthermore, existing co-limitation scenarios continue to emerge, but the interactive effects of iron limitation with other nutrients are relatively unknown.
5. What are the cellular mechanisms for sensing and signaling iron bioavailability and stress? Furthermore, what regulatory mechanisms such as TFs and non-coding RNAs, govern cellular responses to altered iron status?

6. Phytoplankton–bacterial interactions in the context of iron availability remain unclear.

For example, is there competition for iron or is the relationship mutualistic or commensal, particularly with respect to bacterially produced siderophores? For symbiotic relationships such as eukaryotic phytoplankton and endosymbiotic diazotrophs that require high amounts of iron for nitrogenase, are there coordinated responses to iron availability?

Disclosure Statement

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

Acknowledgements

We apologize to all researchers whose work could not be included in this review due to space limitations. This work was supported by the National Science Foundation grants OCE-1756884 and OCE-2224726 (California Current Ecosystem Long-Term Ecological Research), the National Ocean and Atmospheric Administration grant NA19NOS4780181, and the Simons Collaboration on Principles of Microbial Ecosystems (PriME) grant 970820 to A.E.A.

Chapter 1, in full, is a reprint of the material as it appears in Lampe, R.H., Coale, T.H., McQuaid, J.B., and Allen, A.E. in press. *Annual Review in Microbiology* 78. The dissertation author was the primary investigator and author of this paper.

Terms and Definitions:

- Colloidal Iron – nanoparticulate iron species that can pass through a 0.2- or 0.4- μm filter but are not truly dissolved.
- Dissolved Iron – Iron species that can pass through a 0.2- or 0.4- μm filter, thus including both soluble iron species (less than 0.02 μm) and colloidal iron species (nanoparticulate species that are not truly dissolved)
- Particulate Iron – iron within or attached to particles greater 0.2- or 0.4- μm .
- Luxury uptake – the process in which organisms take up more of a nutrient than what is required for maximum growth
- Micronutrient – a chemical element or compound that is required for growth in relatively small quantities
- Siderophore – microbially-produced low molecular weight molecules that have a high-affinity for iron

References

- Allen, A. E., J. Laroche, U. Maheswari, M. Lommer, N. Schauer, P. J. Lopez, G. Finazzi, A. R. Fernie, and C. Bowler. 2008. Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation. *Proc Natl Acad Sci* **105**: 10438-10443.
- Anbar, A. D., and A. H. Knoll. 2002. Proterozoic Ocean Chemistry and Evolution: A Bioinorganic Bridge? *Science* **297**: 1137-1142.
- Anderson, R. F. 2020. GEOTRACES: Accelerating Research on the Marine Biogeochemical Cycles of Trace Elements and Their Isotopes. *Annu Rev Mar Sci* **12**: 49-85.
- Andrew, S. M., C. M. Moreno, K. Plumb, B. Hassanzadeh, L. Gomez-Consarnau, S. N. Smith, O. Schofield, S. Yoshizawa, T. Fujiwara, W. G. Sunda, B. M. Hopkinson, A. N. Septer, and A. Marchetti. 2023. Widespread use of proton-pumping rhodopsin in Antarctic phytoplankton. *Proc Nat Acad Sci USA* **120**: e2307638120.
- Ardyna, M., L. Lacour, S. Sergi, F. d'Ovidio, J.-B. Sallée, M. Rembauville, S. Blain, A. Tagliabue, R. Schlitzer, C. Jeandel, K. R. Arrigo, and H. Claustre. 2019. Hydrothermal vents trigger massive phytoplankton blooms in the Southern Ocean. *Nat Commun* **10**: 2451.
- Armbrust, E. V., J. A. Berges, C. Bowler, B. R. Green, D. Martinez, N. H. Putnam, S. Zhou, A. E. Allen, K. E. Apt, M. Bechner, M. A. Brzezinski, B. K. Chaal, A. Chiovitti, A. K. Davis, M. S. Demarest, J. C. Detter, T. Glavina, D. Goodstein, M. Z. Hadi, U. Hellsten, M. Hildebrand, B. D. Jenkins, J. Jurka, V. V. Kapitonov, N. Kroger, W. W. Lau, T. W. Lane, F. W. Larimer, J. C. Lippmeier, S. Lucas, M. Medina, A. Montsant, M. Obornik, M. S. Parker, B. Palenik, G. J. Pazour, P. M. Richardson, T. A. Rynearson, M. A. Saito, D. C. Schwartz, K. Thamtrakoln, K. Valentin, A. Vardi, F. P. Wilkerson, and D. S. Rokhsar. 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* **306**: 79-86.
- Baker, H. M., B. F. Anderson, and E. N. Baker. 2003. Dealing with iron: Common structural principles in proteins that transport iron and heme. *Proc Nat Acad Sci* **100**: 3579-3583.
- Balk, J., and T. A. Schaedler. 2014. Iron Cofactor Assembly in Plants. *Annu Rev of Plant Biol* **65**: 125-153.
- Behnke, J., Y. Cai, H. Gu, and J. LaRoche. 2023. Short-term response to iron resupply in an iron-limited open ocean diatom reveals rapid decay of iron-responsive transcripts. *PLoS ONE* **18**: e0280827.
- Behnke, J., and J. LaRoche. 2020. Iron uptake proteins in algae and the role of Iron Starvation-Induced Proteins (ISIPs). *Eur J Phycol* **55**: 339-360.
- Behrenfeld, M. J., and A. J. Milligan. 2013. Photophysiological Expressions of Iron Stress in Phytoplankton. *Annu Rev Mar Sci* **5**: 217-246.

Bender, S. J., D. M. Moran, M. R. McIlvin, H. Zheng, J. P. McCrow, J. Badger, G. R. DiTullio, A. E. Allen, and M. A. Saito. 2018. Colony formation in *Phaeocystis antarctica*: connecting molecular mechanisms with iron biogeochemistry. *Biogeosciences* **15**: 4923-4942.

Berthold, D. A., and P. Stenmark. 2003. Membrane-Bound Diiron Carboxylate Proteins. *Annu Rev of Plant Biol* **54**: 497-517.

Bertrand, E. M., J. P. McCrow, A. Moustafa, H. Zheng, J. B. McQuaid, T. O. Delmont, A. F. Post, R. E. Sipler, J. L. Spackeen, K. Xu, D. A. Bronk, D. A. Hutchins, and A. E. Allen. 2015. Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc Nat Acad Sci* **112**: 9938-9943.

Bertrand, E. M., M. A. Saito, J. M. Rose, C. R. Riesselman, M. C. Lohan, A. E. Noble, P. A. Lee, and G. R. DiTullio. 2007. Vitamin B₁₂ and iron colimitation of phytoplankton growth in the Ross Sea. *Limnol Oceanogr* **52**: 1079-1093.

Bidle, K. D., and S. J. Bender. 2008. Iron Starvation and Culture Age Activate Metacaspases and Programmed Cell Death in the Marine Diatom *Thalassiosira pseudonana*. *Eukaryot Cell* **7**: 223-236.

Blaby-Haas, C. E., and S. S. Merchant. 2012. The ins and outs of algal metal transport. *Biochim Biophys Acta Mol Cell Res* **1823**: 1531-1552.

Blaby-Haas, C. E., and S. S. Merchant. 2023. Trace metal nutrition and response to deficiency, p. 167-203. *The Chlamydomonas Sourcebook*. Elsevier.

Botebol, H., E. Lesuisse, R. Šuták, C. Six, J.-C. Lozano, P. Schatt, V. Vergé, A. Kirilovsky, J. Morrissey, T. Léger, J.-M. Camadro, A. Gueneugues, C. Bowler, S. Blain, and F.-Y. Bouget. 2015. Central role for ferritin in the day/night regulation of iron homeostasis in marine phytoplankton. *Proc Natl Acad Sci* **112**: 14652-14657.

Bowler, C., A. Vardi, and A. E. Allen. 2010. Oceanographic and Biogeochemical Insights from Diatom Genomes. *Annu Rev Mar Sci* **2**: 333-365.

Boyd, P. W. 2002. Environmental factors controlling phytoplankton processes in the Southern Ocean. *J Phycol* **38**: 844-861.

Boyd, P. W., M. J. Ellwood, A. Tagliabue, and B. S. Twining. 2017. Biotic and abiotic retention, recycling and remineralization of metals in the ocean. *Nat Geosci* **10**: 167.

Brembu, T., M. Jorstad, P. Winge, K. C. Valle, and A. M. Bones. 2011. Genome-wide profiling of responses to cadmium in the diatom *Phaeodactylum tricornutum*. *Environ Sci Technol* **45**: 7640-7647.

Brown, K. L., K. I. Twing, and D. L. Robertson. 2009. Unraveling the regulation of Nitrogen assimilation in the marine diatom *Thalassiosira pseudonana* (Bacillariophyceae): Diurnal variations in transcript levels for five genes involved in nitrogen assimilation. *J Phycol* **45**: 413-426.

- Browning, T. J., E. P. Achterberg, A. Engel, and E. Mawji. 2021. Manganese co-limitation of phytoplankton growth and major nutrient drawdown in the Southern Ocean. *Nat Commun* **12**: 884.
- Browning, T. J., E. P. Achterberg, I. Rapp, A. Engel, E. M. Bertrand, A. Tagliabue, and C. M. Moore. 2017. Nutrient co-limitation at the boundary of an oceanic gyre. *Nature* **551**: 242-246.
- Brzezinski, M. A., S. B. Baines, W. M. Balch, C. P. Beucher, F. Chai, R. C. Dugdale, J. W. Krause, M. R. Landry, A. Marchi, C. I. Measures, D. M. Nelson, A. E. Parker, A. J. Poulton, K. E. Selph, P. G. Strutton, A. G. Taylor, and B. S. Twining. 2011. Co-limitation of diatoms by iron and silicic acid in the equatorial Pacific. *Deep-Sea Res Pt II* **58**: 493-511.
- Byrne, R. H., and D. R. Kester. 1976. Solubility of hydrous ferric oxide and iron speciation in seawater. *Mar Chem* **4**: 255-274.
- Carradec, Q., E. Pelletier, C. Da Silva, A. Alberti, Y. Seeleuthner, R. Blanc-Mathieu, G. Lima-Mendez, F. Rocha, L. Tirichine, K. Labadie, A. Kirilovsky, A. Bertrand, S. Engelen, M.-A. Madoui, R. Méheust, J. Poulain, S. Romac, D. J. Richter, G. Yoshikawa, C. Dimier, S. Kandels-Lewis, M. Picheral, S. Searson, S. G. Acinas, E. Boss, M. Follows, G. Gorsky, N. Grimsley, L. Karp-Boss, U. Krzic, S. Pesant, E. G. Reynaud, C. Sardet, M. Sieracki, S. Speich, L. Stemmann, D. Velayoudon, J. Weissenbach, O. Jaillon, J.-M. Aury, E. Karsenti, M. B. Sullivan, S. Sunagawa, P. Bork, F. Not, P. Hingamp, J. Raes, L. Guidi, H. Ogata, C. de Vargas, D. Iudicone, C. Bowler, P. Wincker, and C. Tara Oceans. 2018. A global ocean atlas of eukaryotic genes. *Nat Commun* **9**: 373.
- Chappell, P. D., L. P. Whitney, J. R. Wallace, A. I. Darer, S. Jean-Charles, and B. D. Jenkins. 2015. Genetic indicators of iron limitation in wild populations of *Thalassiosira oceanica* from the northeast Pacific Ocean. *ISME J* **9**: 592-602.
- Coale, K. H. 1991. Effects of iron, manganese, copper, and zinc enrichments on productivity and biomass in the subarctic Pacific. *Limnol Oceanogr* **36**: 1851-1864.
- Coale, T. H., E. M. Bertrand, R. H. Lampe, and A. E. Allen. 2022. Molecular Mechanisms Underlying Micronutrient Utilization in Marine Diatoms, p. 567-604. *In* A. Falciatore and T. Mock [eds.], *The Molecular Life of Diatoms*. Springer International Publishing.
- Coale, T. H., M. Moosburner, A. Horák, M. Oborník, K. A. Barbeau, and A. E. Allen. 2019. Reduction-dependent siderophore assimilation in a model pennate diatom. *Proc Natl Acad Sci* **116**: 23609-23617.
- Cohen, N. R., K. A. Ellis, W. G. Burns, R. H. Lampe, N. Schuback, Z. Johnson, S. Sañudo-Wilhelmy, and A. Marchetti. 2017a. Iron and vitamin interactions in marine diatom isolates and natural assemblages of the Northeast Pacific Ocean. *Limnol Oceanogr* **62**: 2076-2096.
- Cohen, N. R., K. A. Ellis, R. H. Lampe, H. M. McNair, B. S. Twining, M. A. Brzezinski, M. T. Maldonado, F. I. Kuzminov, K. Thamtrakoln, C. P. Till, K. W. Bruland, W. G. Sunda, S. Bargu, and A. Marchetti. 2017b. Variations in diatom transcriptional responses to changes in iron availability across ocean provinces. *Front Mar Sci* **4**: 360.

- Cohen, N. R., W. Gong, D. M. Moran, M. R. McIlvin, M. A. Saito, and A. Marchetti. 2018a. Transcriptomic and proteomic responses of the oceanic diatom *Pseudo-nitzschia granii* to iron limitation. *Environ Microbiol* **20**: 3109-3126.
- Cohen, N. R., E. Mann, B. Stemple, S. Raushenberg, J. Jacquot, C. M. Moreno, W. G. Sunda, B. S. Twining, and A. Marchetti. 2018b. Iron storage capacities and associated ferritin gene expression among marine diatoms. *Limnol Oceanogr* **63**: 1677-1691.
- Crawford, D. W., M. S. Lipsen, D. A. Purdie, M. C. Lohan, P. J. Statham, F. A. Whitney, J. N. Putland, W. K. Johnson, N. Sutherland, T. D. Peterson, P. J. Harrison, and C. S. Wong. 2003. Influence of zinc and iron enrichments on phytoplankton growth in the northeastern subarctic Pacific. *Limnol Oceanogr* **48**: 1583-1600.
- Croft, M. T., M. J. Warren, and A. G. Smith. 2006. Algae need their vitamins. *Eukaryot Cell* **5**: 1175-1183.
- Cui, H., H. Ma, Y. Cui, X. Zhu, S. Qin, and R. Li. 2019. Cloning, identification and functional characterization of two cytochrome P450 carotenoids hydroxylases from the diatom *Phaeodactylum tricornutum*. *J Biosci Bioeng* **128**: 755-765.
- D'Ippolito, G., N. Lamari, M. Montresor, G. Romano, A. Cutignano, A. Gerech, G. Cimino, and A. Fontana. 2009. 15S-Lipoxygenase metabolism in the marine diatom *Pseudo-nitzschia delicatissima*. *New Phytol* **183**: 1064-1071.
- Davey, M., and R. J. Geider. 2001. Impact of iron limitation on the photosynthetic apparatus of the diatom *Chaetoceros muelleri* (Bacillariophyceae). *J Phycol* **37**: 987-1000.
- Deng, X., and M. Eriksson. 2007. Two Iron-Responsive Promoter Elements Control Expression of *FOX1* in *Chlamydomonas reinhardtii*. *Eukaryot Cell* **6**: 2163-2167.
- Diaz, J. M., and S. Plummer. 2018. Production of extracellular reactive oxygen species by phytoplankton: past and future directions. *J Plankton Res* **40**: 655-666.
- Ellwood, M. J., and K. A. Hunter. 2000. The incorporation of zinc and iron into the frustule of the marine diatom *Thalassiosira pseudonana*. *Limnol Oceanogr* **45**: 1517-1524.
- Fan, X., H. Qiu, W. Han, Y. Wang, D. Xu, X. Zhang, D. Bhattacharya, and N. Ye. 2020. Phytoplankton pangenome reveals extensive prokaryotic horizontal gene transfer of diverse functions. *Sci Adv* **6**: eaba0111.
- Ferreira, G. C., R. Franco, S. G. Lloyd, I. Moura, J. J. Moura, and B. H. Huynh. 1995. Structure and function of ferrochelatase. *J Bioenerg Biomembr* **27**: 221-229.
- Fitzsimmons, J. N., and T. M. Conway. 2023. Novel Insights into Marine Iron Biogeochemistry from Iron Isotopes. *Annu Rev Mar Sci* **15**: 383-406.
- Fridovich, I. 1998. Oxygen toxicity: a radical explanation. *J Exp Biol* **201**: 1203-1209.

- Geider, R. J., and J. La Roche. 1994. The role of iron in phytoplankton photosynthesis, and the potential for iron-limitation of primary productivity in the sea. *Photosynth Res* **39**: 275-301.
- Geider, R. J., J. La Roche, R. M. Greene, and M. Olaizola. 1993. Response of the photosynthetic apparatus of *Phaeodactylum tricornutum* (Bacillariophyceae) to nitrate, phosphate, or iron starvation. *J Phycol* **29**: 755-766.
- GEOTRACES Intermediate Data Product Group 2023. The GEOTRACES Intermediate Data Product 2021v2 (IDP2021v2). NERC EDS British Oceanographic Data Centre NOC.
- Gerringa, L. J. A., A.-C. Alderkamp, P. Laan, C.-E. Thuróczy, H. J. W. De Baar, M. M. Mills, G. L. van Dijken, H. v. Haren, and K. R. Arrigo. 2012. Iron from melting glaciers fuels the phytoplankton blooms in Amundsen Sea (Southern Ocean): Iron biogeochemistry. *Deep-Sea Res Pt II* **71-76**: 16-31.
- Gilardi, G., and G. Di Nardo. 2017. Heme iron centers in cytochrome P450: structure and catalytic activity. *Rend Lincei* **28**: 159-167.
- Gledhill, M., and K. Buck. 2012. The Organic Complexation of Iron in the Marine Environment: A Review. *Front Microbiol* **3**: 69.
- Goldman, J. A. L., M. J. Schatz, C. T. Berthiaume, S. N. Coesel, M. V. Orellana, and E. V. Armbrust. 2019. Fe limitation decreases transcriptional regulation over the diel cycle in the model diatom *Thalassiosira pseudonana*. *PLoS ONE* **14**: e0222325.
- Gordon, R. M., J. H. Martin, and G. A. Knauer. 1982. Iron in north-east Pacific waters. *Nature* **299**: 611-612.
- Graff van Creveld, S., S. N. Coesel, S. Blaskowski, R. D. Groussman, M. J. Schatz, and E. V. Armbrust. 2023. Divergent functions of two clades of flavodoxin in diatoms mitigate oxidative stress and iron limitation. *eLife* **12**: e84392.
- Groussman, R. D., M. S. Parker, and E. V. Armbrust. 2015. Diversity and Evolutionary History of Iron Metabolism Genes in Diatoms. *PLoS ONE* **10**: e0129081.
- Gruber, N., D. Clement, B. R. Carter, R. A. Feely, S. van Heuven, M. Hoppema, M. Ishii, R. M. Key, A. Kozyr, S. K. Lauvset, C. Lo Monaco, J. T. Mathis, A. Murata, A. Olsen, F. F. Perez, C. L. Sabine, T. Tanhua, and R. Wanninkhof. 2019. The oceanic sink for anthropogenic CO₂ from 1994 to 2007. *Science* **363**: 1193-1199.
- Guérin, N., M. Ciccarella, E. Flamant, P. Frémont, S. Mangenot, B. Istace, B. Noel, C. Belser, L. Bertrand, K. Labadie, C. Cruaud, S. Romac, C. Bachy, M. Gachenot, E. Pelletier, A. Alberti, O. Jaillon, P. Wincker, J.-M. Aury, and Q. Carradec. 2022. Genomic adaptation of the picoeukaryote *Pelagomonas calceolata* to iron-poor oceans revealed by a chromosome-scale genome sequence. *Commun Biol* **5**: 983.
- Halliwell, B., and J. M. C. Gutteridge. 1992. Biologically relevant metal ion-dependent hydroxyl radical generation An update. *FEBS Letters* **307**: 108-112.

- Han, Y., J. Cui, J. Tao, L. Guo, P. Guo, M. Sun, J. Kang, X. Zhang, C. Yan, and S. Li. 2009. CREG inhibits migration of human vascular smooth muscle cells by mediating IGF-II endocytosis. *Exp Cell Res* **315**: 3301-3311.
- Hanikenne, M., S. S. Merchant, and P. Hamel. 2009. Transition metal nutrition: a balance between deficiency and toxicity, p. 333-399. *The Chlamydomonas Sourcebook*. Elsevier.
- Harrison, G. I., and F. M. Morel. 1986. Response of the marine diatom *Thalassiosira weissflogii* to iron stress. *Limnol Oceanogr* **31**: 989-997.
- Hogle, S. L., C. L. Dupont, B. M. Hopkinson, A. L. King, K. N. Buck, K. L. Roe, R. K. Stuart, A. E. Allen, E. L. Mann, Z. I. Johnson, and K. A. Barbeau. 2018. Pervasive iron limitation at subsurface chlorophyll maxima of the California Current. *Proc Nat Acad Sci* **115**: 13300-13305.
- Hopwood, M. J., C. Santana-González, J. Gallego-Urrea, N. Sanchez, E. P. Achterberg, M. V. Ardelan, M. Gledhill, M. González-Dávila, L. Hoffmann, Ø. Leiknes, J. M. Santana-Casiano, T. M. Tsagaraki, and D. Turner. 2020. Fe(II) stability in coastal seawater during experiments in Patagonia, Svalbard, and Gran Canaria. *Biogeosciences* **17**: 1327-1342.
- Hutchins, D. A., and K. W. Bruland. 1998. Iron-limited diatom growth and Si:N uptake ratios in a coastal upwelling regime. *Nature* **393**: 561-564.
- Hutchins, D. A., P. N. Sedwick, G. R. DiTullio, P. W. Boyd, B. Quéguiner, F. B. Griffiths, and C. Crossley. 2001. Control of phytoplankton growth by iron and silicic acid availability in the subantarctic Southern Ocean: Experimental results from the SAZ Project. *J Geophys Res Oceans* **106**: 31559-31572.
- Hutchins, D. A., A. E. Witter, A. Butler, and G. W. Luther. 1999. Competition among marine phytoplankton for different chelated iron species. *Nature* **400**: 858-861.
- Ingall, E. D., J. M. Diaz, A. F. Longo, M. Oakes, L. Finney, S. Vogt, B. Lai, P. L. Yager, B. S. Twining, and J. A. Brandes. 2013. Role of biogenic silica in the removal of iron from the Antarctic seas. *Nat Commun* **4**: 1981.
- Jabre, L. J., A. E. Allen, J. S. P. McCain, J. P. McCrow, N. Tenenbaum, J. L. Spackeen, R. E. Sipler, B. R. Green, D. A. Bronk, D. A. Hutchins, and E. M. Bertrand. 2021. Molecular underpinnings and biogeochemical consequences of enhanced diatom growth in a warming Southern Ocean. *Proc Nat Acad Sci* **118**: e2107238118.
- Kazamia, E., R. Sutak, J. Paz-Yepes, R. G. Dorrell, F. R. J. Vieira, J. Mach, J. Morrissey, S. Leon, F. Lam, E. Pelletier, J.-M. Camadro, C. Bowler, and E. Lesuisse. 2018. Endocytosis-mediated siderophore uptake as a strategy for Fe acquisition in diatoms. *Sci Adv* **4**: eaar4536.
- Keeling, P. J. 2010. The endosymbiotic origin, diversification and fate of plastids. *Philos Trans R Soc Lond B* **365**: 729-748.
- Kolody, B. C., S. R. Smith, L. Z. Allen, J. P. McCrow, A. Moustafa, D. Shi, B. M. Hopkinson, F. M. M. Morel, B. B. Ward, and A. E. Allen. 2022. Nitrogen and Iron Availability Drive

Metabolic Remodeling and Natural Selection of Diverse Phytoplankton during Experimental Upwelling. *mSystems* **7**: e00729-00722.

Kotabova, E., R. Malych, J. J. P. Karlusich, E. Kazamia, M. Eichner, J. Mach, E. Lesuisse, C. Bowler, O. Prášil, and R. Sutak. 2021. Complex Response of the Chlorarachniophyte *Bigelowiella natans* to Iron Availability. *mSystems* **6**: e00738-00720.

Kuma, K., S. Nakabayashi, Y. Suzuki, I. Kudo, and K. Matsunaga. 1992. Photo-reduction of Fe(III) by dissolved organic substances and existence of Fe(II) in seawater during spring blooms. *Mar Chem* **37**: 15-27.

La Roche, J., R. J. Geider, L. M. Graziano, H. Murray, and K. Lewis. 1993. Induction of specific proteins in eukaryotic algae grow under iron-, phosphorus-, or nitrogen-deficient conditions. *J Phycol* **29**: 767-777.

Lampe, R. H., T. H. Coale, K. O. Forsch, L. J. Jabre, S. Kekuewa, E. M. Bertrand, A. Horák, M. Oborník, A. J. Rabines, E. Rowland, H. Zheng, A. J. Andersson, K. A. Barbeau, and A. E. Allen. 2023. Short-term acidification promotes diverse iron acquisition and conservation mechanisms in upwelling-associated phytoplankton. *Nat Commun* **14**: 7215.

Lampe, R. H., G. Hernandez, Y. Y. Lin, and A. Marchetti. 2021. Representative Diatom and Coccolithophore Species Exhibit Divergent Responses throughout Simulated Upwelling Cycles. *mSystems* **6**: e00188-00121.

Lampe, R. H., E. L. Mann, N. R. Cohen, C. P. Till, K. Thamtrakoln, M. A. Brzezinski, K. W. Bruland, B. S. Twining, and A. Marchetti. 2018. Different iron storage strategies among bloom-forming diatoms. *Proc Natl Acad Sci* **115**: E12275-E12284.

Landing, W., and K. Bruland. 1981. The vertical distribution of iron in the northeast Pacific. *Eos, Transactions American Geophysical Union* **62**: 906.

Landing, W. M., and K. W. Bruland. 1987. The contrasting biogeochemistry of iron and manganese in the Pacific Ocean. *Geochim et Cosmochim Acta* **51**: 29-43.

Layer, G., K. Verfu, E. Mahlitz, and D. Jahn. 2002. Oxygen-independent coproporphyrinogen-III oxidase HemN from *Escherichia coli*. *J Biol Chem* **277**: 34136-34142.

Lee, S., B. G. Seok, S.-J. Lee, and S. W. Chung. 2022. Inhibition of mitoNEET attenuates LPS-induced inflammation and oxidative stress. *Cell Death Dis* **13**: 127.

Lelandais, G., I. Scheiber, J. Paz-Yepes, J.-C. Lozano, H. Botebol, J. Pilátová, V. Žárský, T. Léger, P.-L. Blaiseau, C. Bowler, F.-Y. Bouget, J.-M. Camadro, R. Sutak, and E. Lesuisse. 2016. *Ostreococcus tauri* is a new model green alga for studying iron metabolism in eukaryotic phytoplankton. *BMC Genomics* **17**: 319.

Lesuisse, E., P.-L. Blaiseau, A. Dancis, and J.-M. Camadro. 2001. Siderophore uptake and use by the yeast *Saccharomyces cerevisiae*. *Microbiol* **147**: 289-298.

- Lesuisse, E., and P. Labbe. 1989. Reductive and Non-reductive Mechanisms of Iron Assimilation by the Yeast *Saccharomyces cerevisiae*. *Microbiol* **135**: 257-263.
- Lewin, J. C., and R. R. L. Guillard. 1963. Diatoms. *Annu Rev Microbiol* **17**: 373-414.
- Li, T., X. Lin, L. Yu, S. Lin, I. B. Rodriguez, and T.-Y. Ho. 2020. RNA-seq profiling of *Fugacium kawagutii* reveals strong responses in metabolic processes and symbiosis potential to deficiencies of iron and other trace metals. *Sci Total Environ* **705**: 135767.
- Lindqvist, A., and S. Andersson. 2002. Biochemical properties of purified recombinant human β -carotene 15, 15'-monooxygenase. *J Biol Chem* **277**: 23942-23948.
- Lis, H., Y. Shaked, C. Kranzler, N. Keren, and F. M. M. Morel. 2015. Iron bioavailability to phytoplankton: an empirical approach. *ISME J* **9**: 1003-1013.
- Liu, F., M. Gledhill, Q.-G. Tan, K. Zhu, Q. Zhang, P. Salaün, A. Tagliabue, Y. Zhang, D. Weiss, E. P. Achterberg, and Y. Korchev. 2022. Phycosphere pH of unicellular nano- and micro-phytoplankton cells and consequences for iron speciation. *ISME J* **16**: 2329-2336.
- Liu, X., and F. J. Millero. 2002. The solubility of iron in seawater. *Mar Chem* **77**: 43-54.
- Lommer, M., M. Specht, A.-S. Roy, L. Kraemer, R. Andreson, M. A. Gutowska, J. Wolf, S. V. Bergner, M. B. Schilhabel, U. C. Klostermeier, R. G. Beiko, P. Rosenstiel, M. Hippler, and J. LaRoche. 2012. Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol* **13**: R66.
- Lorenzo, M. R., M. Segovia, J. T. Cullen, and M. T. Maldonado. 2020. Particulate trace metal dynamics in response to increased CO₂ and iron availability in a coastal mesocosm experiment. *Biogeosciences* **17**: 757-770.
- Maldonado, M. T., A. E. Allen, J. S. Chong, K. Lin, D. Leus, N. Karpenko, and S. L. Harris. 2006. Copper-dependent iron transport in coastal and oceanic diatoms. *Limnol Oceanogr* **51**: 1729-1743.
- Maldonado, M. T., and N. M. Price. 2001. Reduction and transport of organically bound iron by *Thalassiosira Oceanica* (Bacillariophyceae). *J Phycol* **37**: 298-310.
- Maldonado, M. T., R. F. Strzepek, S. Sander, and P. W. Boyd. 2005. Acquisition of iron bound to strong organic complexes, with different Fe binding groups and photochemical reactivities, by plankton communities in Fe-limited subantarctic waters. *Glob Biogeochem Cycles* **19**, GB4S23.
- Malych, R., P. Stopka, J. Mach, E. Kotabová, O. Prášil, and R. Sutak. 2022. Flow cytometry-based study of model marine microalgal consortia revealed an ecological advantage of siderophore utilization by the dinoflagellate *Amphidinium carterae*. *Comput Struct Biotechnol J* **20**: 287-295.

- Manck, L. E., J. Park, B. J. Tully, A. M. Poire, R. M. Bundy, C. L. Dupont, and K. A. Barbeau. 2022. Petrobactin, a siderophore produced by *Alteromonas*, mediates community iron acquisition in the global ocean. *ISME J* **16**: 358-369.
- Maniscalco, M. A., M. A. Brzezinski, R. H. Lampe, N. R. Cohen, H. M. McNair, K. A. Ellis, M. Brown, C. P. Till, B. S. Twining, K. W. Bruland, A. Marchetti, and K. Thamatrakoln. 2022. Diminished carbon and nitrate assimilation drive changes in diatom elemental stoichiometry independent of silicification in an iron-limited assemblage. *ISME Commun* **2**: 57.
- Marchetti, A., and M. T. Maldonado. 2016. Iron, p. 233-279. *In* M. A. Borowitzka, J. Beardall and J. A. Raven [eds.], *The Physiology of Microalgae*. Springer International Publishing.
- Marchetti, A., M. S. Parker, L. P. Moccia, E. O. Lin, A. L. Arrieta, F. Ribalet, M. E. P. Murphy, M. T. Maldonado, and E. V. Armbrust. 2009. Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature* **457**: 467-470.
- Marchetti, A., D. M. Schruth, C. A. Durkin, M. S. Parker, R. B. Kodner, C. T. Berthiaume, R. Morales, A. E. Allen, and E. V. Armbrust. 2012. Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci* **109**: E317-E325.
- Marchetti, A., D. E. Varela, V. P. Lance, V. P. Lance, M. Palmucci, M. Giordano, and E. Virginia Armbrust. 2010. Iron and silicic acid effects on phytoplankton productivity, diversity, and chemical composition in the central equatorial Pacific Ocean. *Limnol Oceanogr* **55**: 11-29.
- Martin, J. H., and S. E. Fitzwater. 1988. Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic. *Nature* **331**: 341-343.
- Martin, J. H., R. M. Gordon, S. Fitzwater, and W. W. Broenkow. 1989. Vertex: phytoplankton/iron studies in the Gulf of Alaska. *Deep-Sea Res Oceanogr, A* **36**: 649-680.
- Martinez-Gomez, N. C., and D. M. Downs. 2008. ThiC Is an [Fe-S] Cluster Protein That Requires AdoMet To Generate the 4-Amino-5-hydroxymethyl-2-methylpyrimidine Moiety in Thiamin Synthesis. *Biochem* **47**: 9054-9056.
- Mausz, M. A., M. Segovia, A. Larsen, S. A. Berger, J. K. Egge, and G. Pohnert. 2020. High CO₂ concentration and iron availability determine the metabolic inventory in an *Emiliania huxleyi*-dominated phytoplankton community. *Environ Microbiol* **22**: 3863-3882.
- McCain, J. S. P., A. Tagliabue, E. Susko, E. P. Achterberg, A. E. Allen, and E. M. Bertrand. 2021. Cellular costs underpin micronutrient limitation in phytoplankton. *Sci Adv* **7**: eabg6501.
- McQuaid, J. B., A. B. Kustka, M. Oborník, A. Horák, J. P. McCrow, B. J. Karas, H. Zheng, T. Kindeberg, A. J. Andersson, K. A. Barbeau, and A. E. Allen. 2018. Carbonate-sensitive phytotransferrin controls high-affinity iron uptake in diatoms. *Nature* **555**: 534.

- Mitchell, C. J., A. Shawki, T. Ganz, E. Nemeth, and B. Mackenzie. 2014. Functional properties of human ferroportin, a cellular iron exporter reactive also with cobalt and zinc. *Am J Physiol Cell Physiol* **306**: C450-C459.
- Moffett, J. W., and R. M. Boiteau. 2024. Metal Organic Complexation in Seawater: Historical Background and Future Directions. *Annu Rev Mar Sci* **16**: 577-599.
- Moore, J. K., S. C. Doney, D. M. Glover, and I. Y. Fung. 2001. Iron cycling and nutrient-limitation patterns in surface waters of the World Ocean. *Deep-Sea Res Pt II* **49**: 463-507.
- Morel, F. M. M., A. B. Kustka, and Y. Shaked. 2008. The role of unchelated Fe in the iron nutrition of phytoplankton. *Limnol Oceanogr* **53**: 400-404.
- Morrissey, J., R. Sutak, J. Paz-Yepes, A. Tanaka, A. Moustafa, A. Veluchamy, Y. Thomas, H. Botebol, F.-Y. Bouget, Jeffrey B. McQuaid, L. Tirichine, Andrew E. Allen, E. Lesuisse, and C. Bowler. 2015. A Novel Protein, Ubiquitous in Marine Phytoplankton, Concentrates Iron at the Cell Surface and Facilitates Uptake. *Curr Biol* **25**: 364-371.
- Murik, O., L. Tirichine, J. Prihoda, Y. Thomas, W. L. Araújo, A. E. Allen, A. R. Fernie, and C. Bowler. 2019. Downregulation of mitochondrial alternative oxidase affects chloroplast function, redox status and stress response in a marine diatom. *New Phytol* **221**: 1303-1316.
- Nishino, T., and K. Okamoto. 2000. The role of the [2Fe-2S] cluster centers in xanthine oxidoreductase. *J Inorg Biochem* **82**: 43-49.
- Nuester, J., S. Vogt, and B. S. Twining. 2012. Localization of iron within centric diatoms of the genus *Thalassiosira*. *J Phycol* **48**: 626-634.
- Nunn, B. L., J. F. Faux, A. A. Hippmann, M. T. Maldonado, H. R. Harvey, D. R. Goodlett, P. W. Boyd, and R. F. Strzepek. 2013. Diatom Proteomics Reveals Unique Acclimation Strategies to Mitigate Fe Limitation. *PLoS ONE* **8**: e75653.
- Oborník, M., and B. R. Green. 2005. Mosaic origin of the heme biosynthesis pathway in photosynthetic eukaryotes. *Mol Biol Evol* **22**: 2343-2353.
- Paz, Y., E. Shimoni, M. Weiss, and U. Pick. 2007. Effects of Iron Deficiency on Iron Binding and Internalization into Acidic Vacuoles in *Dunaliella salina*. *Plant Physiol* **144**: 1407-1415.
- Peers, G., and N. M. Price. 2006. Copper-containing plastocyanin used for electron transport by an oceanic diatom. *Nature* **441**: 341-344.
- Peers, G., S.-A. Quesnel, and N. M. Price. 2005. Copper requirements for iron acquisition and growth of coastal and oceanic diatoms. *Limnol Oceanogr* **50**: 1149-1158.
- Pierella Karlusich, J. J., A. F. Lodeyro, and N. Carrillo. 2014. The long goodbye: the rise and fall of flavodoxin during plant evolution. *J Exp Bot* **65**: 5161-5178.

- Provasoli, L. 1958. Nutrition and Ecology of Protozoa and Algae. *Annu Rev Microbiol* **12**: 279-308.
- Puig, S., L. Ramos-Alonso, A. M. Romero, and M. T. Martínez-Pastor. 2017. The elemental role of iron in DNA synthesis and repair. *Metallomics* **9**: 1483-1500.
- Raiswell, R., L. G. Benning, M. Tranter, and S. Tulaczyk. 2008. Bioavailable iron in the Southern Ocean: the significance of the iceberg conveyor belt. *Geochemical Trans* **9**: 7.
- Raven, J. A. 1988. The iron and molybdenum use efficiencies of plant growth with different energy, carbon and nitrogen sources. *New Phytol* **109**: 279-287.
- Raven, J.A. 1990. Predictions of Mn and Fe use efficiencies of phototrophic growth as a function of light availability for growth and of C assimilation pathway. *New Phytol* **116**: 1-18.
- Raven, J. A., and J. Beardall. 2017. Consequences of the genotypic loss of mitochondrial Complex I in dinoflagellates and of phenotypic regulation of Complex I content in other photosynthetic organisms. *J Exp Bot* **68**: 2683-2692.
- Raven, J. A., M. C. W. Evans, and R. E. Korb. 1999. The role of trace metals in photosynthetic electron transport in O₂-evolving organisms. *Photosynth Res* **60**: 111-150.
- Read, A. D., R. E. T. Bentley, S. L. Archer, and K. J. Dunham-Snary. 2021. Mitochondrial iron-sulfur clusters: Structure, function, and an emerging role in vascular biology. *Redox Biol* **47**: 102164.
- Rijkenberg, M. J. A., R. Middag, P. Laan, L. J. A. Gerringa, H. M. van Aken, V. Schoemann, J. T. M. de Jong, and H. J. W. de Baar. 2014. The Distribution of Dissolved Iron in the West Atlantic Ocean. *PLoS ONE* **9**: e101323.
- Roncel, M., D. Kirilovsky, F. Guerrero, A. Serrano, and J. M. Ortega. 2012. Photosynthetic cytochrome *c*₅₅₀. *Biochim Biophys Acta Bioenerg* **1817**: 1152-1163.
- Ryan-Keogh, T. J., and S. J. Thomalla. 2020. Deriving a Proxy for Iron Limitation From Chlorophyll Fluorescence on Buoyancy Gliders. *Front Mar Sci* **7**, 275.
- Santana-Casiano, J. M., M. González-Dávila, and F. J. Millero. 2005. Oxidation of Nanomolar Levels of Fe(II) with Oxygen in Natural Waters. *Environ Sci Technol* **39**: 2073-2079.
- Scheiber, I. F., J. Pilátová, R. Malych, E. Kotabova, M. Krijt, D. Vyoral, J. Mach, T. Léger, J.-M. Camadro, O. Prášil, E. Lesuisse, and R. Sutak. 2019. Copper and iron metabolism in *Ostreococcus tauri* – the role of phytoferritin, plastocyanin and a chloroplast copper-transporting ATPase. *Metallomics* **11**: 1657-1666.
- Segovia, M., M. R. Lorenzo, M. T. Maldonado, A. Larsen, S. A. Berger, T. M. Tsagaraki, F. J. Lázaro, C. Iñiguez, C. García-Gómez, A. Palma, M. A. Mausz, F. J. L. Gordillo, J. A. Fernández, J. L. Ray, and J. K. Egge. 2017. Iron availability modulates the effects of future CO₂ levels within the marine planktonic food web. *Mar Ecol Prog Ser* **565**: 17-33.

- Shaked, Y., K. N. Buck, T. Mellett, and M. T. Maldonado. 2020. Insights into the bioavailability of oceanic dissolved Fe from phytoplankton uptake kinetics. *ISME J* **14**: 1182-1193.
- Shaked, Y., A. B. Kustka, and F. M. M. Morel. 2005. A general kinetic model for iron acquisition by eukaryotic phytoplankton. *Limnol Oceanogr* **50**: 872-882.
- Shi, D., Y. Xu, B. M. Hopkinson, and F. M. M. Morel. 2010. Effect of Ocean Acidification on Iron Availability to Marine Phytoplankton. *Science* **327**: 676-679.
- Shire, D. M. 2022. Investigating mechanisms of trace metal metabolism and subcellular protein localization in marine phytoplankton. Rutgers The State University of New Jersey, Graduate School.
- Shire, D. M., and A. B. Kustka. 2015. Luxury uptake, iron storage and ferritin abundance in *Prochlorococcus marinus* (Synechococcales) strain MED4. *Phycologia* **54**: 398-406.
- Smith, S. R., C. L. Dupont, J. K. McCarthy, J. T. Broddrick, M. Oborník, A. Horák, Z. Füßy, J. Cihlář, S. Kleessen, H. Zheng, J. P. McCrow, K. K. Hixson, W. L. Araújo, A. Nunes-Nesi, A. Fernie, Z. Nikoloski, B. O. Palsson, and A. E. Allen. 2019. Evolution and regulation of nitrogen flux through compartmentalized metabolic networks in a marine diatom. *Nat Commun* **10**: 4552.
- Smith, S. R., J. T. F. Gillard, A. B. Kustka, J. P. McCrow, J. H. Badger, H. Zheng, A. M. New, C. L. Dupont, T. Obata, A. R. Fernie, and A. E. Allen. 2016. Transcriptional Orchestration of the Global Cellular Response of a Model Pennate Diatom to Diel Light Cycling under Iron Limitation. *PLoS Genet* **12**: e1006490.
- Soria-Dengg, S., and U. Horstmann. 1995. Ferrioxamines B and E as iron sources for the marine diatom *Phaeodactylum tricornutum*. *Mar Ecol Prog Ser* **127**: 269-277.
- Strauss, J., L. Deng, S. Gao, A. Toseland, C. Bachy, C. Zhang, A. Kirkham, A. Hopes, R. Utting, E. F. Joest, A. Tagliabue, C. Löw, A. Z. Worden, G. Nagel, and T. Mock. 2023. Plastid-localized xanthorhodopsin increases diatom biomass and ecosystem productivity in iron-limited surface oceans. *Nat Microbiol* **8**: 2050-2066.
- Strickland, J. D. 1965. Phytoplankton and marine primary production. *Annu Rev Microbiol* **19**: 127-162.
- Strzepek, R. F., P. W. Boyd, and W. G. Sunda. 2019. Photosynthetic adaptation to low iron, light, and temperature in Southern Ocean phytoplankton. *Proc Nat Acad Sci* **116**: 4388-4393.
- Strzepek, R. F., and P. J. Harrison. 2004. Photosynthetic architecture differs in coastal and oceanic diatoms. *Nature* **431**: 689-692.
- Strzepek, R. F., M. T. Maldonado, K. A. Hunter, R. D. Frew, and P. W. Boyd. 2011. Adaptive strategies by Southern Ocean phytoplankton to lessen iron limitation: Uptake of organically complexed iron and reduced cellular iron requirements. *Limnol Oceanogr* **56**: 1983-2002.

- Sugie, K., H. Endo, K. Suzuki, J. Nishioka, H. Kiyosawa, and T. Yoshimura. 2013. Synergistic effects of pCO₂ and iron availability on nutrient consumption ratio of the Bering Sea phytoplankton community. *Biogeosciences* **10**: 6309-6321.
- Sunda, W. G., and S. A. Huntsman. 1995. Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Mar Chem* **50**: 189-206.
- Sunda, W. G., and A. Marchetti. 2024. Proton-pumping rhodopsins promote the growth and survival of phytoplankton in a highly variable ocean. *ISME J* **18**: wrac079.
- Sutak, R., H. Botebol, P.-L. Blaiseau, T. Léger, F.-Y. Bouget, J.-M. Camadro, and E. Lesuisse. 2012. A Comparative Study of Iron Uptake Mechanisms in Marine Microalgae: Iron Binding at the Cell Surface Is a Critical Step. *Plant Physiol* **160**: 2271-2284.
- Sutak, R., J.-M. Camadro, and E. Lesuisse. 2020. Iron Uptake Mechanisms in Marine Phytoplankton. *Front Microbiol* **11**, 566691.
- Sutak, R., J. Šlapeta, M. San Roman, J.-M. Camadro, and E. Lesuisse. 2010. Nonreductive Iron Uptake Mechanism in the Marine Alveolate *Chromera velia*. *Plant Physiol* **154**: 991-1000.
- Tagliabue, A., L. Bopp, O. Aumont, and K. R. Arrigo. 2009. Influence of light and temperature on the marine iron cycle: From theoretical to global modeling. *Glob Biogeochem Cycles* **23**: GB2017.
- Tagliabue, A., A. R. Bowie, P. W. Boyd, K. N. Buck, K. S. Johnson, and M. A. Saito. 2017. The integral role of iron in ocean biogeochemistry. *Nature* **543**: 51-59.
- Tagliabue, A., K. N. Buck, L. E. Sofen, B. S. Twining, O. Aumont, P. W. Boyd, S. Caprara, W. B. Homoky, R. Johnson, D. König, D. C. Ohnemus, B. Sohst, and P. Sedwick. 2023. Authigenic mineral phases as a driver of the upper-ocean iron cycle. *Nature* **620**: 104-109.
- Tagliabue, A., T. Mtshali, O. Aumont, A. R. Bowie, M. B. Klunder, A. N. Roychoudhury, and S. Swart. 2012. A global compilation of dissolved iron measurements: focus on distributions and processes in the Southern Ocean. *Biogeosciences* **9**: 2333-2349.
- Tan, M. H., S. R. Smith, K. K. Hixson, J. Tan, J. K. McCarthy, A. B. Kustka, and A. E. Allen. 2020. The Importance of Protein Phosphorylation for Signaling and Metabolism in Response to Diel Light Cycling and Nutrient Availability in a Marine Diatom. *Biology* **9**: 155.
- Teng, L., X. Fan, D. R. Nelson, W. Han, X. Zhang, D. Xu, H. Renault, G. V. Markov, and N. Ye. 2019. Diversity and evolution of cytochromes P450 in stramenopiles. *Planta* **249**: 647-661.
- Trimborn, S., S. Thoms, T. Brenneis, J. P. Heiden, S. Beszteri, and K. Bischof. 2017. Two Southern Ocean diatoms are more sensitive to ocean acidification and changes in irradiance than the prymnesiophyte *Phaeocystis antarctica*. *Physiol Plant* **160**: 155-170.

- Turnšek, J., J. K. Brunson, M. d. P. M. Viedma, T. J. Deerinck, A. Horák, M. Oborník, V. A. Bielinski, and A. E. Allen. 2021. Proximity proteomics in a marine diatom reveals a putative cell surface-to-chloroplast iron trafficking pathway. *eLife* **10**: e52770.
- Twining, B. S., and S. B. Baines. 2022. Luxury iron uptake and storage in pennate diatoms from the equatorial Pacific Ocean. *Metallomics* **14**: mfac035.
- Vanoni, M. A., and B. Curti. 1999. Glutamate synthase: a complex iron-sulfur flavoprotein. *Cell Mol Life Sci* **55**: 617-638.
- Völker, C., and D. A. Wolf-Gladrow. 1999. Physical limits on iron uptake mediated by siderophores or surface reductases. *Mar Chem* **65**: 227-244.
- Wade, J., D. J. Byrne, C. J. Ballentine, and H. Drakesmith. 2021. Temporal variation of planetary iron as a driver of evolution. *Proc Nat Acad Sci* **118**: e2109865118.
- Ward, D. M., and J. Kaplan. 2012. Ferroportin-mediated iron transport: Expression and regulation. *Biochim Biophys Acta Mol Cell Res* **1823**: 1426-1433.
- Wu, M., J. S. P. McCain, E. Rowland, R. Middag, M. Sandgren, A. E. Allen, and E. M. Bertrand. 2019. Manganese and iron deficiency in Southern Ocean *Phaeocystis antarctica* populations revealed through taxon-specific protein indicators. *Nat Commun* **10**: 3582.
- Yoshinaga, R., M. Niwa-Kubota, H. Matsui, and Y. Matsuda. 2014. Characterization of iron-responsive promoters in the marine diatom *Phaeodactylum tricornutum*. *Mar Genomics* **16**: 55-62.

RELATED RESOURCES

- Blaby-Haas, C.E., and S.S. Merchant. 2017. Regulating cellular trace metal economy in algae. *Curr Opin Plant Biol* **39**: 88-96.
- Hutchins, D.A. and P.W. Boyd. 2016. Marine phytoplankton and the changing ocean iron cycle. *Nat Clim Change* **6**: 1072-1079
- Gao, X., C. Bowler, and E. Kazamia. 2021. Iron metabolism strategies in diatoms. *J Exp Bot* **72**: 2165-80.
- Twining B.S. and S.B. Baines. 2013. The Trace Metal Composition of Marine Phytoplankton. *Annu Rev Mar Sci* **5**: 191-215.

CHAPTER 2 - SHORT-TERM ACIDIFICATION PROMOTES DIVERSE IRON ACQUISITION AND CONSERVATION MECHANISMS IN UPWELLING-ASSOCIATED PHYTOPLANKTON

Robert H. Lampe, Tyler H. Coale, Kiefer O. Forsch, Loay J. Jabre, Samuel Kekuwa, Erin M. Bertrand, Aleš Horák, Miroslav Oborník, Ariel J. Rabines, Elden Rowland, Hong Zheng, Andreas J. Andersson, Katherine A. Barbeau, and Andrew E. Allen

Abstract

Coastal upwelling regions are among the most productive marine ecosystems but may be threatened by amplified ocean acidification. Increased acidification is hypothesized to reduce iron bioavailability for phytoplankton thereby expanding iron limitation and impacting primary production. Here we show from community to molecular levels that phytoplankton in an upwelling region respond to short-term acidification exposure with iron uptake pathways and strategies that reduce cellular iron demand. A combined physiological and multi-omics approach was applied to trace metal clean incubations that introduced 1200 ppm CO₂ for up to four days. Although variable, molecular-level responses indicate a prioritization of iron uptake pathways that are less hindered by acidification and reductions in iron utilization. Growth, nutrient uptake, and community compositions remained largely unaffected suggesting that these mechanisms may confer short-term resistance to acidification; however, we speculate that cellular iron demand is only temporarily satisfied, and longer-term acidification exposure without increased iron inputs may result in increased iron stress.

Introduction

Present-day atmospheric carbon dioxide (CO₂) concentrations of over 400 ppm have not been observed in the past 800,000 years and are increasing at an average rate that exceeds any observed in the past 20,000 years (Joos and Spahni 2008; Keeling et al. 2005; Tripathi et al.

2009). Approximately 30% of this released CO₂ is absorbed in the surface ocean resulting in a shift in the carbonic acid system and decline in seawater pH known as ocean acidification (Gruber et al. 2019).

Coastal upwelling regions such as the California Current System (CCS) account for a significant proportion of marine primary production and exhibit some of the largest pH variability in the ocean due to the natural acidity of upwelled waters (Capone and Hutchins 2013; Leinweber and Gruber 2013). Some coastal upwelling areas are also expected to experience accelerated or amplified ocean acidification (Capone and Hutchins 2013; Hauri et al. 2013). In the CCS for example, the rate of acidification may be double that of the global average resulting in increased overall exposure to higher acidity for organisms, particularly during periods of strong upwelling (Capone and Hutchins 2013; Feely et al. 2008; Osborne et al. 2020).

Phytoplankton, especially diatoms, drive the high levels of primary production in upwelling regions that in turn support highly productive fisheries (Capone and Hutchins 2013); however, their vulnerability to acidification remains uncertain. Most phytoplankton species appear to exhibit relatively little response to acidification although there is variability among taxa (Mackey et al. 2015). Field-based experiments with diatom communities suggest that effects from acidification are more frequently positive than negative and coastal communities are less affected than oceanic communities; however, shifts in community composition may also occur (Bach and Taucher 2019).

Further uncertainty in phytoplankton responses lies in the relationships between ocean acidification and iron bioavailability. Iron is a critical micronutrient for phytoplankton growth that is often limiting in the ocean including coastal upwelling regions such as the California Current and Peru/Humboldt Current Systems (Bruland et al. 2001; Hutchins et al. 2002; Moore

et al. 2013; Till et al. 2019), and its availability in a future more acidic ocean remains controversial (Hutchins and Boyd 2016). Reductions in bioavailable iron in response to acidification have been demonstrated in laboratory settings and are attributed to two mechanisms: (1) under acidic conditions, iron is less likely to disassociate from certain organic ligands to its more bioavailable form, labile dissolved inorganic iron (Fe') (Shi et al. 2010), and (2) the algal inorganic iron uptake protein, phytoferritin (pTF), is dependent on carbonate ion concentrations ($[\text{CO}_3^{2-}]$) that decline under acidic conditions (McQuaid et al. 2018). In the California Current system, $[\text{CO}_3^{2-}]$ is estimated to have declined 35% during the 20th century (Osborne et al. 2020).

The laboratory studies demonstrating these impacts in bioavailable iron may not have captured important dynamics including biotic interactions with siderophore-producing bacteria, a significantly more complex iron pool, and dynamic nutrient availability (Boyd et al. 2015). Further acidification experiments with natural seawater and a single diatom species, *Thalassiosira weissflogii*, did not produce significant differences in iron uptake rates indicating the presence of natural iron-binding ligands that are largely unaffected by pH (Shi et al. 2010), although reductions in Fe' in the natural environment from increased complexation have also been observed (Avendaño et al. 2016; Gledhill et al. 1998). Other experiments and modelling studies suggest that iron bioavailability may instead increase in response to acidification due to increased solubility of Fe(III) (Gledhill et al. 2015; Stockdale et al. 2016; Zhu et al. 2021), increased oxidative lifetimes of highly bioavailable Fe(II) (Breitbarth et al. 2010; Capone and Hutchins 2013), and enhanced dissolution of particulate iron (Lorenzo et al. 2020).

Previous field-based studies examining the effects of ocean acidification on phytoplankton communities under iron stress or limitation, namely within the iron-limited

subarctic North Pacific and Southern Ocean, have shown conflicting results. Some experiments have shown that diatom communities are negatively impacted with reduced abundances, growth rates, and photosynthetic efficiencies (Endo et al. 2015; Pausch et al. 2022; Sugie et al. 2013; Trimborn et al. 2017). Other experiments observed no changes in growth, macronutrient uptake, and community composition (Endo et al. 2013; Yoshimura et al. 2014; Yoshimura et al. 2013). One experiment in the Southern Ocean observed no differences in growth or macronutrient uptake in the whole phytoplankton community; however, centric diatoms appeared to be favoured over pennate diatoms indicating variable responses among more specific taxonomic groups that may not be evident if examining the entire community (Feng et al. 2010). In contrast to these other experiments, Hopkinson et al. (2010) made pH adjustments to communities in the Gulf of Alaska resulting in very modest increases to growth and photosynthetic efficiency in conjunction with downregulation of certain photosynthetic proteins. Other ocean acidification studies have been conducted in regions with relatively high ambient iron concentrations, some of which used artificially high levels of strong iron-binding ligands to induce iron stress (Chen et al. 2004; Hoppe et al. 2013; Mausz et al. 2020; Segovia et al. 2017); however, these ligands also increase the solubility of iron leading to increased overall dissolved iron concentrations (Borer et al. 2005).

Within upwelling regions, previous experiments also indicate a lack of change in response to acidification with diatom-dominated phytoplankton assemblages, but acidification was not assessed along with iron status even though the phytoplankton communities in these regions often exhibit high iron demand and limitation (Bruland et al. 2001; Hutchins et al. 2002; Till et al. 2019). In the CCS, several experiments showed no discernible differences in growth, primary productivity, and carbon-to-nitrogen ratios (Tortell et al. 2002; Tortell et al. 2000;

Tortell et al. 1997); however, responses were not evaluated with CO₂ levels exceeding 800 ppm to exceed modern day surface pH variability in upwelling systems (Supplementary Tables 1 and 2). Only Osma, et al. (2020) examined responses up to 1,600 ppm CO₂ in the Peru/Humboldt Current System, but also did not observe differences in macronutrient drawdown, chlorophyll *a* concentrations, or diatom cell abundances.

High pH variability in coastal upwelling ecosystems has led to speculation that the native microorganisms in these regions may have the metabolic flexibility to accommodate relatively large pH changes. As evidenced by the response of higher trophic levels (mussels, gastropods, and planktonic copepods) in the Peru/Humboldt Current System, examining ocean acidification relative to local conditions rather than the global mean provides an improved understanding of organisms' tolerance (Vargas et al. 2017). As a result, phytoplankton responses in coastal upwelling regions may be different than those observed in other areas such as chronically iron-limited (high-nitrate low-chlorophyll) regions with low pH variability (Capone and Hutchins 2013; Joint et al. 2011).

Furthermore, the aforementioned field-based experiments employed physiological assessments of the whole phytoplankton community that are incapable of discriminating changes only occurring within a specific taxonomic group. At times, taxon-specific pigment analyses and microscopic cell counts were employed showing a lack of change in abundances for more specific taxonomic groups or species, but not offering a window into taxon-specific physiological adjustments (Endo et al. 2015; Endo et al. 2013; Hopkinson et al. 2010). As previously mentioned, centric and pennate diatoms exhibited differing trends in their relative abundances in an experiment in the Southern Ocean (Feng et al. 2010) suggesting more taxonomically-specific resolution is needed to evaluate responses to ocean acidification.

Transcriptomic and proteomic analyses can provide this greater resolution while also providing insight into the short-term molecular-level changes that may precede physiological changes or differences in community structure. Moreover, molecular-level knowledge of phytoplankton responses to acidification is limited to a small number of taxa under laboratory conditions without considering iron bioavailability or using organisms obtained from upwelling areas (Beszteri et al. 2018; Jones et al. 2013; Valenzuela et al. 2018).

Thus, to evaluate potential impacts of acidification on phytoplankton in an upwelling region, ocean acidification was simulated with natural phytoplankton communities in the CCS. Trace metal clean techniques were used throughout enabling interrogation of acidification-driven effects on iron availability without unnatural changes to the iron pool. Air-CO₂ gas mixtures up to 1200 ppm CO₂ were applied, which exceeds both the projected global mean atmospheric CO₂ concentration for the year 2100 under the highest emission Shared Socio-Economic Pathway (SSP5-8.5) and most reported surface seawater pCO₂ and pH variability in the CCS (Supplementary Tables 1-3). With a combined physiological and multi-omic approach, these experiments enable probing of mixed phytoplankton assemblages' sensitivity to short-term acidification exposure from the whole community to taxonomically-specific molecular levels while considering the high importance of iron bioavailability.

Results

Experimental set-up, initial conditions, and assessments of iron limitation

Four incubation experiments were conducted in the CCS. Experiments 1-3 occurred within an upwelling filament, represented by newly upwelled waters at Experiment 1 that progressively aged by Experiment 3, and used communities from the near-surface (Fig. 2.1a and

Supplemental Fig. 2.1, Table 2.1). Experiment 4 serves as a farther offshore site with an initial community from a subsurface chlorophyll maximum layer (SCML) at 60 m (Supplemental Fig. 2.1).

Initial pH at Experiment 1 was lowest (7.89 ± 0.00) as upwelled waters are naturally acidic and then became less acidic (8.06 ± 0.01 by Experiments 3 and 4) from outgassing and primary production (Supplemental Fig. 2.2 and Supplemental Table 2.3)(Capone and Hutchins 2013; Feely et al. 2008; Fiechter et al. 2014; Turi et al. 2014). In each experiment, triplicate bottles were supplemented with chelexed macronutrients (Table 2.1) and bubbled with HEPA-filtered air balanced with 400, 800, or 1200 ppm CO₂. Samples were then collected at one-day intervals that varied among experiments to attempt to avoid macronutrient depletion with an upper limit of 96 h to also avoid increased bottle effects: 48 and 72 h for experiment 1, 24 and 48 h for experiment 2, and 48 and 96 h for experiments 3 and 4 (Table 2.1). By the time of sampling, the incubations had an approximate pH of 8.0 ± 0.0 , 7.8 ± 0.0 , and 7.6 ± 0.0 (Supplemental Fig. 2.2) and pCO₂ concentrations of 413 ± 21 , 800 ± 34 , and 1197 ± 44 μ atm for each respective treatment with some variability in pCO₂ concentrations among experiments (Supplemental Table 2.3).

As anticipated for the region, initial dissolved iron concentrations ([dFe]) were below 1 nmol L⁻¹ (Fig. 2.1b)(Bruland et al. 2001); only a small fraction of this dissolved iron exists as inorganic Fe(III) with most bound to organic ligands (Bundy et al. 2014). Labile particulate iron concentrations were highest in the vicinity of Experiment 1 (> 10 nmol L⁻¹), and decreased to generally less than 1.0 nmol L⁻¹ offshore near Experiments 3 and 4 (Supplemental Fig. 2.3).

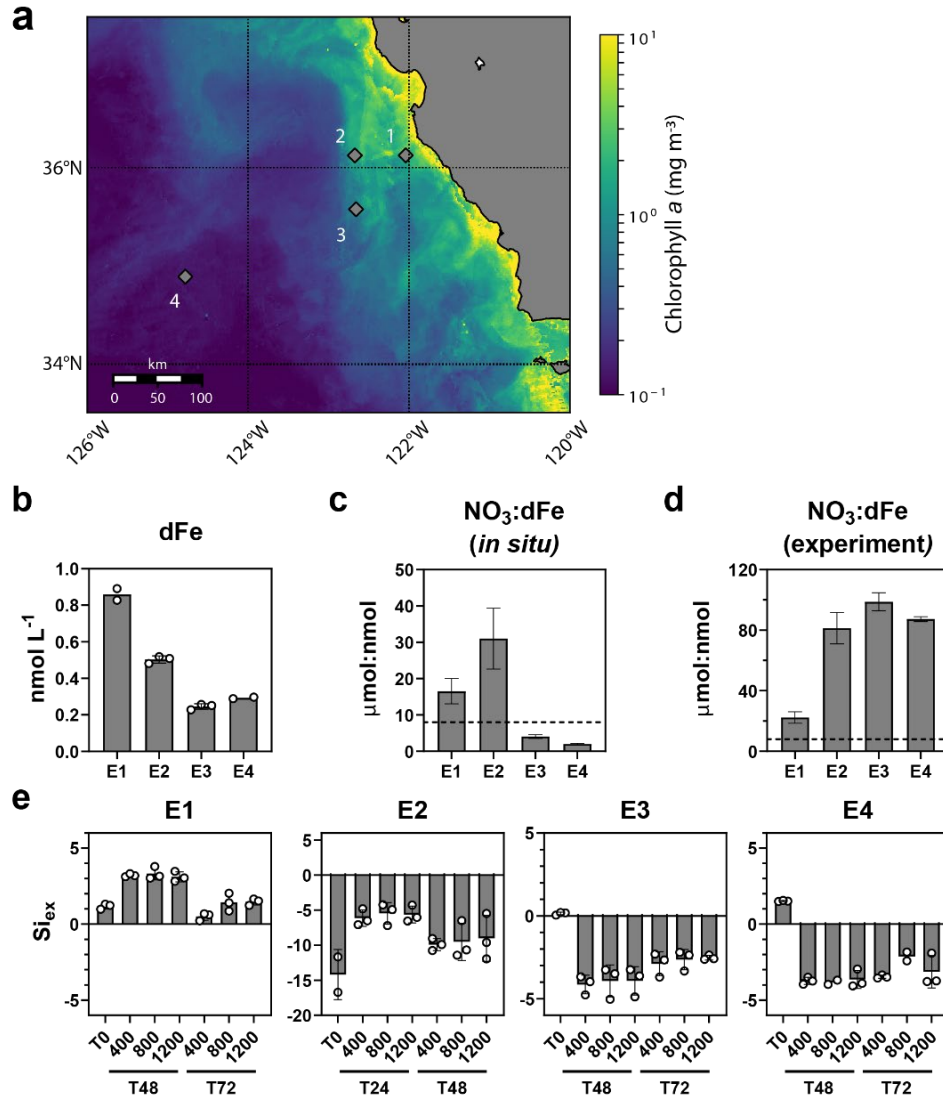


Figure 2.1: Overview of the initial conditions for the four experiments (E1-E4). (a) Map of experiment locations with satellite-derived chlorophyll *a*. The chlorophyll *a* data was obtained from the NOAA Coastwatch Browser as a 14-day average centred on August 22, 2020 which falls between Experiments 2 and 3 (Table 1). (b) Dissolved iron concentrations (dFe, nmol L⁻¹) (c, d) Nitrate (NO₃, μmol L⁻¹) to dissolved iron (dFe, nmol L⁻¹) ratios (c) before and (d) after macronutrient addition at the start of each experiment. The horizontal dashed line indicates the threshold for potential iron limitation (8 μmol:nmol)(Hogle et al. 2018; King and Barbeau 2011). Ratios were derived and standard deviations were propagated from the average values of field replicates presented in Figs. 2.1a and 2.2a (e) Si_{ex} proxy for diatom iron limitation within each experimental treatment. Initial time point (T0) values reflect the Si_{ex} value following macronutrient addition. Negative values indicate iron stress or limitation due to preferential uptake of silicic acid (H₄SiO₄) relative to nitrate (NO₃)(Hogle et al. 2018; King and Barbeau 2011). Error bars represent the standard deviation of the mean ($n = 3$ biologically independent samples or field replicate samples for T0). For dissolved iron concentrations at Experiments 1 and 4 and macronutrient data at T0 of Experiment 2, $n = 2$.

Table 2.1: Experiment information. All experiments were conducted during the same cruise in August, 2019. Macronutrients were added at the concentrations specified. Subsampling was performed for all measurements at the time points listed (hours) in addition to the initial time point (T0). Chlorophyll a concentrations were measured every 24 h.

Exp.	Date	Latitude & Longitude	NO ₃ & H ₄ SiO ₄ added (μmol L ⁻¹)	PO ₄ added (μmol L ⁻¹)	Time points (hours)
1	Aug. 13	36° 7.398 N 122° 2.475 W	5	0.3125	48, 72
2	Aug. 19	36° 7.333 N 122° 40.441 W	20	1.25	24, 48
3	Aug. 24	35° 34.823 N 122° 39.551 W	25	1.5625	48, 96
4	Aug. 30	34° 53.818 N 124° 46.905 W	25	1.5625	48, 96

With these iron concentrations, the addition of macronutrients created the potential for iron stress or limitation, particularly at Experiments 3 and 4 with Experiment 2 already showing signs of Fe limitation *in situ*. Ratios of nitrate to dissolved iron concentrations (NO₃:dFe) greater than eight (μmol:nmol) suggest the potential for iron limitation (Hogle et al. 2018), and the initial conditions of all experiments exceeded this threshold following the addition of NO₃ (Fig. 2.1c, d). These ratios are also in line with those that naturally occur in the region (Forsch et al. 2023).

Si_{ex} is an additional biogeochemical proxy that serves as an indicator of Fe limitation in diatoms (Hogle et al. 2018); as Fe-limited diatoms preferentially uptake silicic acid (H₄SiO₄) relative to NO₃ (Hutchins and Bruland 1998), lower than expected H₄SiO₄ concentrations compared to NO₃ concentrations indicates the presence of Fe limitation. Si_{ex} reports the difference between H₄SiO₄ and NO₃ concentrations (*Methods*); therefore, negative values indicate greater H₄SiO₄ uptake due to iron stress in diatoms. At Experiment 2, negative values

were observed from the initial timepoint (T0) throughout the experiment (Fig. 2.1e). Following macronutrient addition, negative values were also observed in Experiments 3 and 4 while values remained positive in Experiment 1 (Fig. 2.1e).

Parallel incubations with unamended control treatments and iron addition treatments (+5 nmol L⁻¹) were performed to further evaluate the presence of iron stress during the experiments. Following 24 hours of incubation at Experiment 2 and 48 hours at Experiment 3, chlorophyll *a* concentrations significantly increased with the addition of iron indicating the presence of iron limitation in these communities by the first sampling time points (Supplemental Fig. 2.4). At Experiment 4, nitrate was initially depleted indicating no iron limitation in the initial community (Fig. 2.2 and Supplemental Fig. 2.5); however, in incubations initiated 24 hours later while following a Lagrangian drifter, significant increases in chlorophyll *a* were observed only with the addition of both iron and nitrate suggesting co-limitation in this community. As nitrate was added without iron in the community subjected to increased acidification, it was likely iron-limited. Considering low (< 1 μmol L⁻¹) nitrate concentrations after 72 hours in Experiment 1 (Fig. 2.2), the phytoplankton community there was not iron-limited, although it may have encountered some degree of iron stress.

Whole-community physiological responses

Despite differences in pH between treatments at all timepoints (Supplemental Fig. 2.2), few significant differences ($P < 0.05$) were observed in chlorophyll *a* concentrations, particulate carbon-to-nitrogen ratios, and macronutrient concentrations. Short-term iron uptake rates were normalized to both chlorophyll *a* and particulate organic carbon (POC) with both inorganic iron (FeCl₃) and iron bound to the organic ligand desferrioxamine B (Fe-DFB)(Fig. 2.2 and

Supplemental Figs. 2.6, 2.7). Although chlorophyll content varies in response to environmental conditions (Wang et al. 2009), light and temperature were consistent during the experiments resulting in comparable chlorophyll-normalized rates that may better reflect photoautotrophic biomass compared to POC. Chlorophyll *a* has often also been used to normalize Fe uptake rates in phytoplankton communities as POC includes heterotrophic biomass (Hoppe et al. 2013; Marchetti and Maldonado 2016; Trimborn et al. 2015).

When normalized to chlorophyll *a*, maximum inorganic Fe uptake rates were significantly higher in the 1200 ppm when compared to the 400 ppm treatment at the final time point for Experiment 3 ($P = 0.016$) indicating an increase in cell surface transporters (Marchetti and Maldonado 2016). All other Fe uptake rates were not significantly different among treatments at each time point and experiment including when normalized to POC. The measurable uptake of Fe-DFB, particularly at Experiments 3 and 4 where rates were similar to those of iron-limited populations and isolates (Maldonado and Price 1999; Maldonado and Price 2001), further indicates iron stress as phytoplankton normally only acquire significant amounts of Fe-DFB when iron-limited (Marchetti and Maldonado 2016).

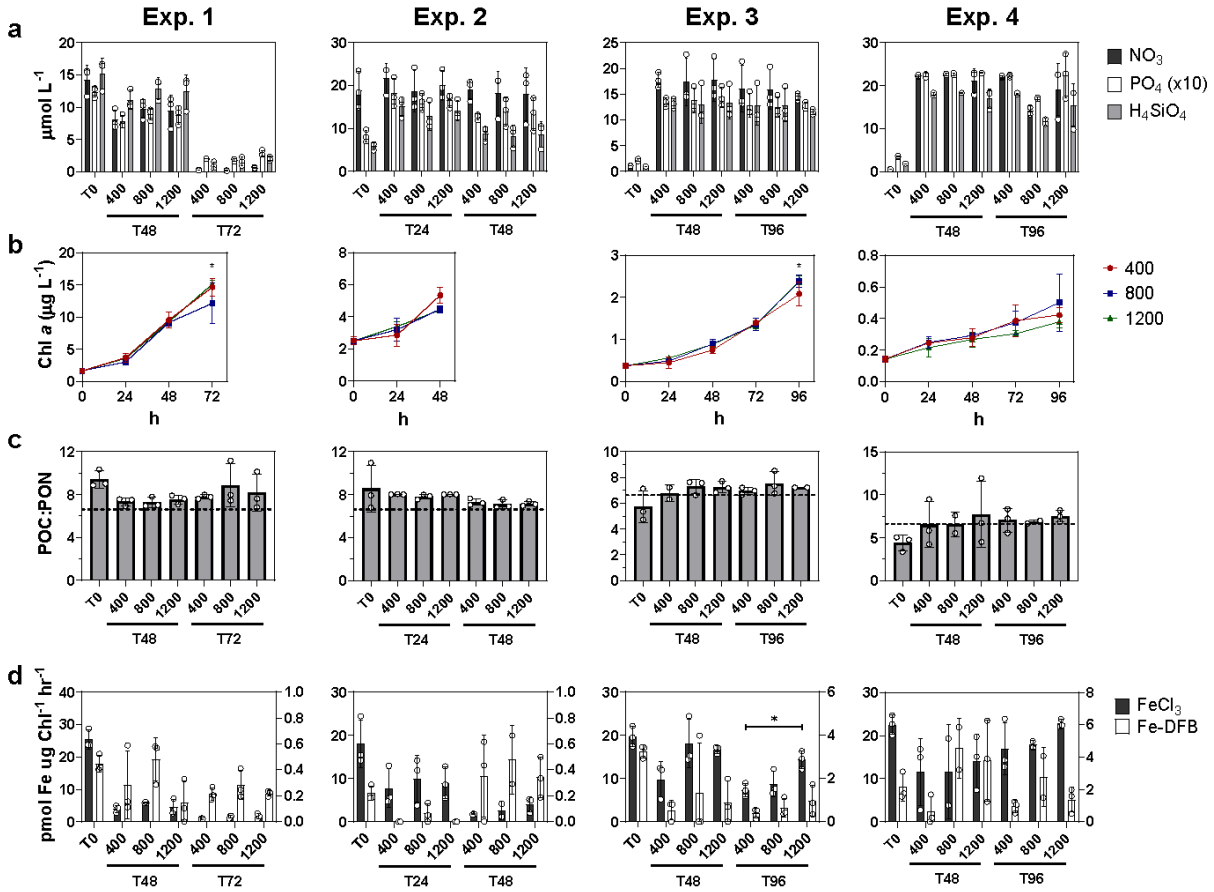


Figure 2.2: Physiological measurements associated with the incubations. (a) Nitrate (NO_3 , black), phosphate ($\text{PO}_4 \times 10$, white), and silicate (H_4SiO_4 , gray) concentrations ($\mu\text{mol L}^{-1}$). The initial (T0) concentrations are those before amending at the concentrations shown in Table 2.1. (b) Chlorophyll *a* concentrations ($\mu\text{mol L}^{-1}$) from every 24 h interval for the 400 (red), 800 (blue), and 1200 ppm CO_2 treatments (green). At Experiment 1, the 800 ppm and 1200 ppm treatments are significantly different ($P = 0.026$) after 72 h. At Experiment 3, the 400 ppm treatment is significantly different from both the 800 ppm ($P = 0.018$) and 1200 ppm treatments ($P = 0.010$) after 96 h. (c) Ratios of particulate organic carbon (POC) to particulate organic nitrogen (PON). The expected Redfield ratio (106:16) is denoted with horizontal dashed line. (d) Short-term iron-59 uptake rates normalized to chlorophyll *a* ($\text{pmol Fe } \mu\text{g Chl}^{-1} \text{ hr}^{-1}$) for inorganic iron (FeCl_3 , black, left y-axis) and organically-complexed iron as ferrioxamine B (Fe-DFB, white, right y-axis). At Experiment 3, the FeCl_3 uptake rates are significantly different between the 400 and 1200 ppm treatments ($P = 0.016$). Both pCO_2 treatments and time points in hours are denoted on the x-axis in a, c, and d. Error bars represent the standard deviation of the mean ($n = 3$ biologically independent samples or field replicate samples for T0 except the Experiment 4 800 ppm samples where $n = 2$). Significant differences ($P < 0.05$, Two-way ANOVA with Tukey's HSD test) are denoted with an asterisk (*).

Community growth rates (μ) estimated from chlorophyll *a* concentrations were also not significantly different between the experiments (Supplemental Fig. 2.8). In Experiment 1, a relatively rapid average growth rate of $0.87 \pm 0.10 \text{ d}^{-1}$ suggests that members of these communities may have performed multiple cell divisions leading to a large increase in chlorophyll *a* and drawdown of NO_3 to $<1 \text{ } \mu\text{mol L}^{-1}$. Multiple divisions also may have potentially occurred at Experiment 3 ($\mu = 0.46 \pm 0.04 \text{ d}^{-1}$) and approximately one division may have occurred before the termination of Experiment 4 ($\mu = 0.27 \pm 0.04 \text{ d}^{-1}$) as both experiments were four days long. Collectively, these rates suggest that the total exposure to acidification exceeds that of the community doubling time. In the case of Experiment 2, the shorter duration of the experiment (48 h) and lower average growth rate ($\mu = 0.32 \pm 0.08 \text{ d}^{-1}$) suggests that the community doubling time exceeded the duration of exposure. This lower growth rate was likely driven by iron-limitation in the initial community.

Microbial community composition

Eukaryotic and prokaryotic community compositions throughout the ocean acidification experiments were inferred via 18S and 16S ribosomal RNA (rRNA) amplicon sequencing (Fig. 2.3). Within eukaryotic phytoplankton rRNA, Experiments 1-3 were dominated by centric diatoms, particularly the large chain-forming genera *Chaetoceros*, *Eucampia*, and *Thalassiosira*, with relatively high abundances in Experiment 4 as well (Fig. 2.3a, Supplemental Fig. 2.9). Pennate diatoms were also present throughout and largely consisted of *Pseudo-nitzschia* with a notable presence of *Navicula* in Experiment 2 (Supplemental Fig. 2.9). *Chaetoceros*, *Thalassiosira*, and *Pseudo-nitzschia* are also among the most abundant diatoms in both the region and the ocean (James et al. 2022; Malviya et al. 2016). Experiment 3 had a high

abundance of the prasinophyte *Ostreococcus* (Bathycoccaceae) and the pelagophyte *Pelagomonas calceolata* (Pelagomonadales), both of which are cosmopolitan picoeukaryotes (Demir-Hilton et al. 2011; Worden et al. 2012). *Pelagomonas calceolata* further dominated Experiment 4 in line with their previously observed high abundances in SCMLs (Dupont et al. 2015).

Alphaproteobacteria and Gammaproteobacteria, were prevalent in all four experiments as observed in the global ocean (Fig. 2.3a)(Sunagawa et al. 2015). In particular, these groups were dominated by the ubiquitous *Rhodobacterales* order (Buchan et al. 2005) and SAR86 clade (Hoarfrost et al. 2020)(Supplemental Fig. S.10). Cyanobacteria were less than 0.5% of the prokaryotic communities in Experiments 1 and 2 further indicating that the phytoplankton communities were dominated by diatoms. Experiment 3 had a notable presence of *Synechococcus* (14% of 16S rRNA), and Experiment 4 was dominated by *Prochlorococcus* (71% of 16S rRNA) indicating a higher presence of cyanobacteria in both of those experiments.

Like the physiological measurements, increased CO₂ largely did not alter the eukaryotic or prokaryotic community compositions. Alpha diversity expressed as the Shannon Index was not significantly different among CO₂ treatments in each experiment except for prokaryotes in Experiment 1 where diversity was higher in the 400 ppm treatment (Fig. 2.3b). Community dissimilarity (Bray-Curtis) was driven more by the distinct initial communities and change in the communities over time with no significant differences among CO₂ treatments for both prokaryotes and eukaryotes in each experiment (Fig. 2.3c and Supplemental Fig. 2.11, PERMANOVA, $P > 0.05$).

Differential abundances of specific taxa as rRNA-derived amplicon sequence variants (ASVs) were evaluated with DESeq2 among the treatments at the final timepoints (Fig. 2.3d and

Supplemental Fig. 2.12). Between the 400 ppm and 1200 ppm treatments, DESeq2 estimated that fewer than 1.5% of the eukaryotic ASVs and no more than a single prokaryotic ASV within each experiment were differentially abundant. Five 18S ASVs with significantly decreased abundances at 1200 ppm were haptophytes belonging to the Prymnesiophyceae class or Isochrysidales order without more detailed taxonomic resolution. These taxonomic groups contain calcifying algae that may be more susceptible to acidification as the precipitation of calcium carbonate becomes less energetically favourable, although a wide range of responses for haptophytes has previously been observed and some species are not calcifying (Dutkiewicz et al. 2015; Edvardsen et al. 2000; Mackey et al. 2015). Three additional differentially abundant 18S ASVs were metazoan sequences including a hydrozoan and a copepod, which likely stems from uneven distributions of metazoans between treatments rather than an acidification-driven response.

Differential expression of transcripts and proteins among CO₂ treatments

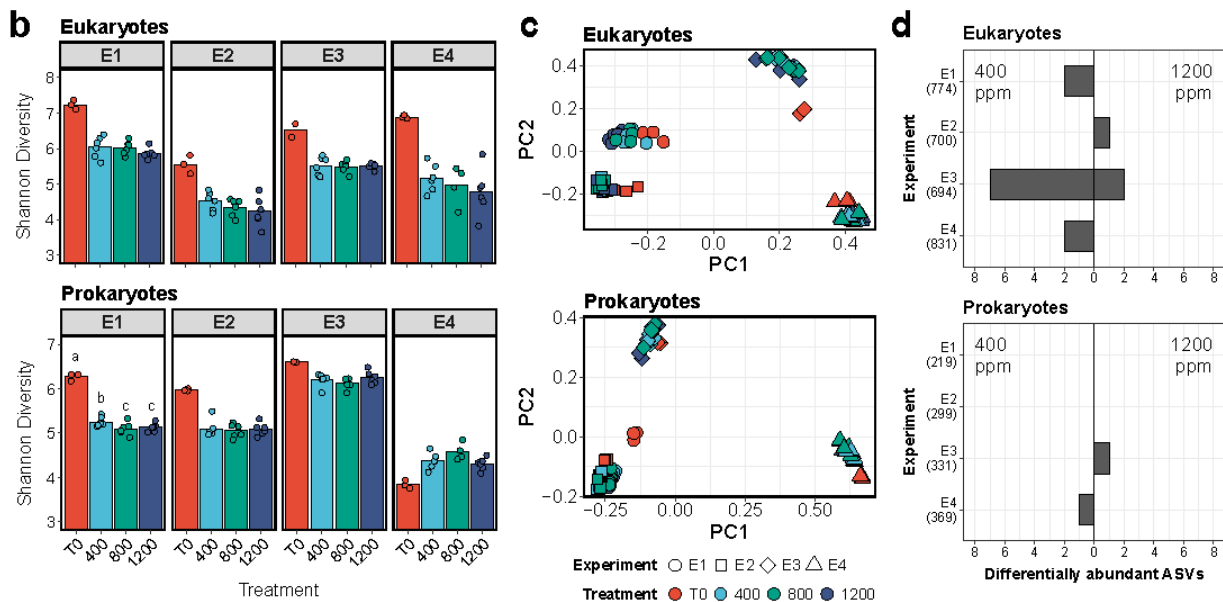
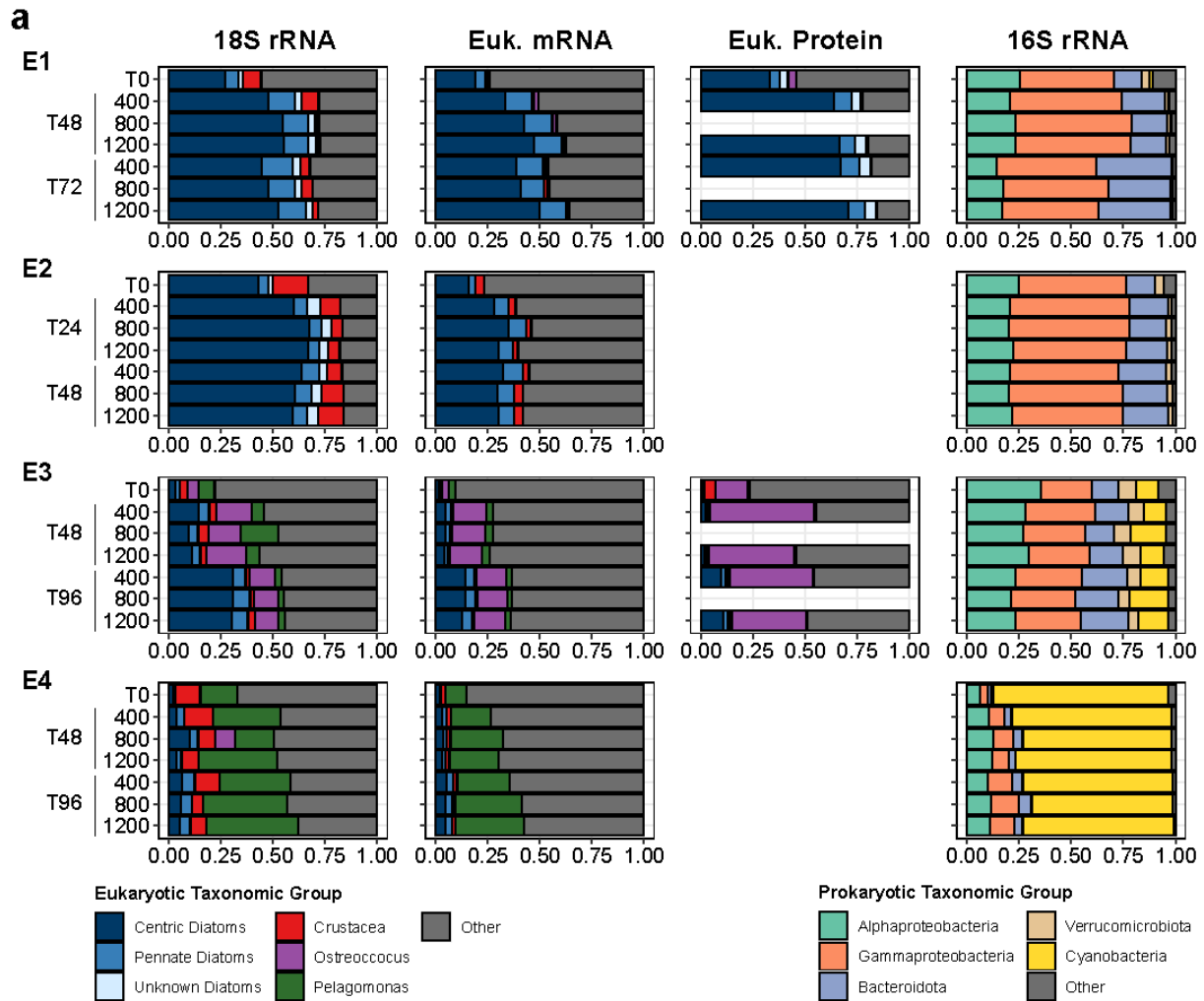
Dominant eukaryotic phytoplankton groups identified from 18S rRNA were abundant in corresponding poly(A)-selected metatranscriptomes and metaproteomes (Fig. 2.3a). Between 23.8% and 48.2% of orthologous gene groups, or orthogroups, were unique to the detected protein sequences in each experiment's metatranscriptome assembly highlighting the diverse functional repertoires within each community (Supplemental Fig. 2.13) in addition to their significant differences in microbial community composition (Fig. 2.3c). In each experiment, differential expression of KEGG-annotated transcripts was examined between CO₂ treatments within these taxonomic groups. The same was conducted for protein abundances in Experiments 1 and 3. Often differentially expressed transcripts were not detected as proteins (Supplemental

2.4), likely due to comparatively lower resolution with mass spectrometry, but some consistent patterns were observed.

Transcriptional changes were generally lower at 800 versus 400 ppm compared to 1200 versus 400 ppm indicating that transcriptional responses are dependent on CO₂ conditions (Fig. 2.4a and Supplemental Table 2.5); protein abundances were not examined for the 800 ppm treatments. Dominant taxonomic groups also displayed differing levels of responsiveness with centric diatoms being more responsive than pennate diatoms in terms of both transcripts and proteins. Further variability was observed in the transcripts or proteins that were differentially abundant within taxonomic groups across experiments and time points.

As acidification is hypothesized to induce iron stress, three indicators of iron stress, the iron starvation-induced genes *ISIP1*, *ISIP2A/pTF*, and *ISIP3*, were examined in both centric and pennate diatoms (Fig. 2.4b and Supplemental Fig. 2.14). *ISIP1* is linked to endocytosis of siderophore-bound iron (Kazamia et al. 2018). *ISIP2A*, or phytotransferrin (pTF), is an inorganic iron uptake protein with a direct dependence on carbonate ions (McQuaid et al. 2018). As previously mentioned, this carbonate ion dependence may result in negative impacts on iron uptake as carbonate ions decline under acidic conditions. *ISIP3* does not have a known function although it has a ferritin-like domain further suggesting a role in iron metabolism (Behnke and LaRoche 2020).

Figure 2.3: Microbial community composition and diversity for each experiment (E1-4) from both 18S and 16S rRNA amplicons. (a) The average relative abundances of major taxonomic groups from 18S rRNA, poly(A)-selected mRNA (metatranscriptome), proteins (metaproteome), and 16S rRNA within each time point and treatment. The average abundances from the metatranscriptomes were derived from the relative abundances of eukaryotic-assigned transcripts per million normalized reads (TPM). The average abundances from the metaproteomes were derived from the average relative abundances of total sum ion intensities for all eukaryotic peptides identified. (b) Alpha diversity expressed as the Shannon Index within each sample for eukaryotic rRNA (top) and prokaryotic rRNA (bottom). Bars indicate the average Shannon Index. For prokaryotic ASVs at Experiment 1, the initial conditions (T0) are significantly different than all other treatments (Kruskal-Wallis test, Benjamini & Hochberg adjusted $P = 0.037$). The 400 ppm treatment is also significantly different from both the 800 ppm (Benjamini & Hochberg adjusted $P = 0.037$) and 1200 ppm (Benjamini & Hochberg adjusted $P = 0.045$) treatments. $n = 3$ biologically independent samples or field replicate samples for T0 except the Experiment 3 T0 and Experiment 4 800 ppm samples where $n = 2$. (c) PCoA plots of Bray-Curtis dissimilarity for eukaryotic rRNA (top) and prokaryotic rRNA (bottom). Experiments and treatments are denoted by color and shape respectively as described in the legend. (d) The number of significantly differential abundant amplicon sequence variants (ASVs) at each experiment determined by DESeq2 (Wald test) between the 400 ppm (left) and 1200 ppm (right) treatments at the final time point for each experiment (Benjamini & Hochberg adjusted $P < 0.05$). Eukaryotic ASVs are shown on top and prokaryotic ASVs are shown on the bottom plot. The total number of ASVs for each experiment is shown in parentheses next to the experiment number of the y-axis.



These genes are often highly responsive to iron status in diatoms (Behnke and LaRoche 2020; Cohen et al. 2017; Smith et al. 2016), and high gene expression relative to others in each sample would not be expected unless the cells were experiencing iron stress or limitation. In Experiments 1 and 2, all three genes were in the top 1% of expressed transcripts by the final time point in all CO₂ treatments indicating extremely high expression, iron stress, and iron demand even though macronutrients were depleted at the end of Experiment 1 (Fig. 2.4b and Supplemental Fig. 2.14). In Experiments 3 and 4, transcript expression was within the top 5% of expressed genes also indicating iron stress. Importantly, none of the *ISIP* transcripts were significantly differentially expressed between CO₂ treatments at each time point ($P > 0.05$) suggesting no relative increase or decrease in iron stress as a function of CO₂. These patterns were also consistent between centric and pennate diatoms (Fig. 2.4b and Supplemental Fig. 2.14).

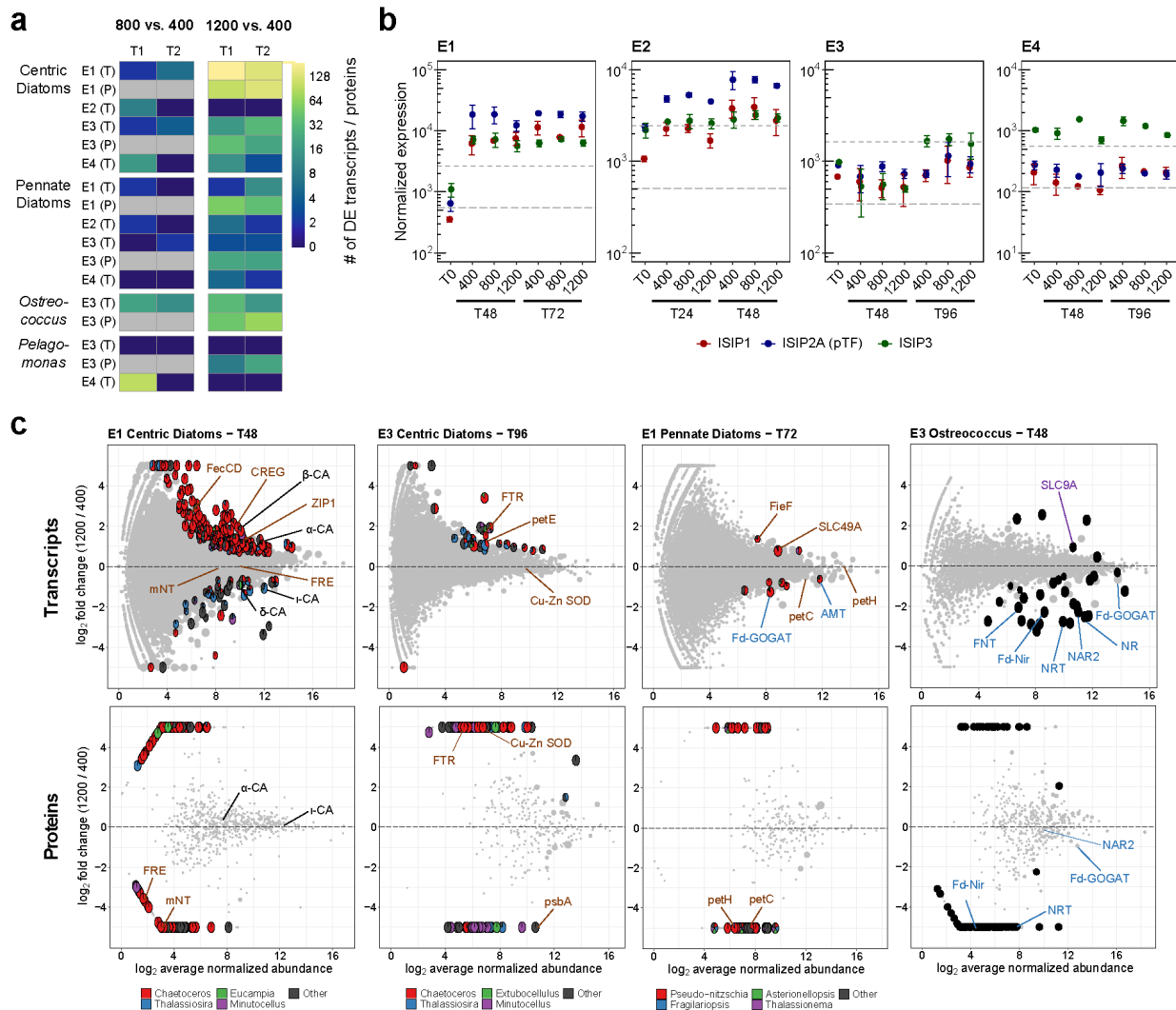
ISIP proteins were not universally detected; however, in Experiments 1 and 3, ISIP3 from centric diatoms was in the top 5% of detected proteins at the final time points further suggesting the development of iron stress in these experiments (Supplemental Fig. 2.15). Other ISIP proteins approached the top 5% in both centric and pennate diatoms. Protein abundances were also not significantly different between CO₂ treatments. The exceptions are for ISIP1 and ISIP2A in centric diatoms in Experiment 3 at 48 h where they were not detected in the 1200 ppm treatment; however, both were detected after 96 h where abundances were not significantly different (Supplemental Fig. 2.15).

Nutrient stress in Experiments 1 and 3 was further evaluated by examining cellular resource allocations to ribosomal and photosynthetic proteins (Supplemental Fig. 2.16) (McCain et al. 2022). At Experiment 1, both centric and pennate diatoms displayed significant reductions

in their ribosomal protein mass fractions from 48 to 72 h, likely reflecting the onset of nitrogen limitation after 72 h ($P < 0.01$, Supplemental Fig. 2.16). At the earlier timepoint, 48 h, pennate diatoms displayed moderate increases in their ribosomal protein mass fractions under higher CO₂ ($P = 0.11$) indicating an increase in ribosomal investment from increased acidity. At Experiment 3, a different response was observed in centric diatoms and *Ostreococcus* with both significantly reducing their photosynthetic protein mass fractions after the initial time point ($P < 0.01$, Supplemental Fig. 2.16). This shift is likely a result of increased iron stress as observed in diatoms within the Southern Ocean (McCain et al. 2022). A lack of change to the ribosomal mass fraction in comparison to pennate diatoms at Experiment 1 further highlights the potential differences in responses among taxonomic groups or the interactive effects of ocean acidification under increased Fe stress.

The total number of significantly differentially expressed transcripts was variable across the experiments with the most found at Experiments 1 and 3 where the exposure to acidification relative to the community growth rates were the longest (Supplemental Fig. 2.8, Supplemental Tables 2.4 and 2.5). Although total differential expression of transcripts within each taxonomic group was less than 1%, those that were differentially expressed highlight the reconfiguration of iron, nitrogen, and carbon uptake proteins as well as cellular iron requirements to increased CO₂. Many other differentially expressed genes are of unknown function as anticipated for eukaryotic phytoplankton (Blaby-Haas and Merchant 2019).

Figure 2.4: Molecular-level responses to increased pCO₂. (a) Heatmap of the number of significantly differentially expressed transcripts (T) or proteins (P) for dominant taxonomic groups at each experiment (DESeq2 Wald test, $P < 0.05$, log₂ scale) in the 800 ppm vs 400 ppm treatments (left) and 1200 ppm vs 400 ppm treatments (right) at each time point. Gray boxes indicate time points where there is no corresponding proteome. (b) Normalized transcript expression of iron starvation-induced proteins (ISIPs) in centric diatoms (log₁₀ scale): *ISIP1* (red), *ISIP2A* (phytotransferrin or *pTF*, blue), *ISIP3* (green). The horizontal dashed lines indicate the averaged 95th and 99th percentile for total transcript abundance in each experiment. Error bars represent the standard deviation of the mean ($n = 3$ biologically independent samples or field replicate samples for T0 except the Experiment 3 T0 and Experiment 4 800 ppm samples where $n = 2$). (c) MA plots for taxonomic groups, experiments (E), and time points (T), where significant differential expression for transcripts (top row) or proteins (bottom row) of interest between the 1200 and 400 ppm treatments was observed. Differentially expressed genes or proteins (DESeq2 Wald test, Benjamini & Hochberg adjusted $P < 0.05$) are shown as pies representing the proportion of expression for dominant genera. Pie size is proportional to P value with larger pies denoting a smaller P value. Black circles rather than pies are shown for *Ostreococcus* as expression is already examined at the genus level. Gray points are genes that were not significantly differentially abundance (DESeq2 Wald test, Benjamini & Hochberg adjusted $P > 0.05$). Data points with log₂ fold changes less than -5 or greater than 5 are plotted at -5 and 5 respectively.



Under higher CO₂, increased expression of genes related to iron uptake or transport that are distinct from *ISIP2a/pTF* was observed (Fig. 2.4c). At times, several of these genes have been observed to be upregulated by diatoms under iron stress in the laboratory and field (Allen et al. 2008; Cohen et al. 2017; Lommer et al. 2012). This response includes an iron permease (*FTR1*). Diatom FTR1 proteins are related to those found in green algae and yeast where they have been shown to mediate inorganic ferric iron uptake like *pTF* (Supplemental Fig. 2.17)(Behnke and LaRoche 2020; Blaby-Haas and Merchant 2012). *FTR1* transcripts and proteins significantly increased in centric diatoms at 1200 ppm CO₂ in Experiment 3 indicating an increase in cell surface transporters. This increase in transporters likely resulted in the significantly higher maximum uptake rates of inorganic iron observed here (Fig. 2.2d)(Harrison and Morel 1986).

Two putative ferrous iron transporters displayed increased transcript expression in Experiment 1 under increased CO₂: one in the ZIP family (ZIP1) in centric diatoms and one in the CDF family (FieF-like) in pennate diatoms (Fig. 2.4c). Diatom homologs of the *ZIP1* gene are related to ZIP-family genes in green algae, and have been found to be upregulated under Fe stress for diatoms in the both the field and laboratory (Allen et al. 2008; Cohen et al. 2017; Lampe et al. 2018). The CDF transporter found here is from a distinct clade than one previously described to be upregulated under Fe stress in the model diatom *P. tricornutum* (Supplemental Fig. 2.18) and is related to green algal CDF transporters that have been shown to have Mn transport capability (Ibuot et al. 2020).

ZIP and CDF-family transporters can transport multiple divalent metal cations (Blaby-Haas and Merchant 2012); therefore, if localized to the cell surface, they may serve as low affinity Fe(II) permeases to take advantage of potentially increased Fe(II)(Breitbarth et al. 2010;

Coale et al. 2022). Direct Fe(II) uptake has been demonstrated in eukaryotic microalgae including diatoms supporting this role(Sutak et al. 2012). ZIP and CDF-family transporters may also be involved with intracellular transport or release of stored iron potentially indicating an increased reliance on iron recycling or storage (Blaby-Haas and Merchant 2012; Kustka et al. 2007; Lampe et al. 2018).

Putative siderophore-bound iron uptake genes, an FecCD-domain containing gene in centric diatoms and SLC49 family transporters in pennate diatoms also displayed increased transcript expression at 1200 ppm CO₂ in Experiment 1 (Fig. 2.4c). These genes are poorly characterized in diatoms but have been linked to siderophore-based Fe uptake in other organisms. The FecCD domain has been characterized as part of a bacterial operon for a ferric citrate ABC transporter in *E. coli* (Staudenmaier et al. 1989). Similar proteins are present in several centric and pennate diatoms and were likely obtained via horizontal gene transfer from bacteria (Supplemental Fig. 2.19). Diatom SLC49-family transporters are similar to those found in metazoans where they have been functionally characterized (Supplemental Fig. 2.20)(Khan and Quigley 2013). These proteins are related to heme transport and may be responsible for heme uptake as observed in diatom cultures and eukaryotic phytoplankton in the natural environment (Hutchins et al. 1999; Khan and Quigley 2013).

The increased expression of these transporters is consistent with the expected decrease in bioavailable free inorganic ferric iron that may result from acidification. At Experiment 1, increased transporter expression implies an increased reliance on siderophore-bound iron or ferrous iron that may increase unlike ferric iron. At Experiment 3, the response was instead an increase of a ferric iron permease (FTR1) that does not have a carbonate-dependence like phytoferritin. Observed as both transcripts and proteins, this increase suggests that there were

a greater number of cell surface transporters for inorganic ferric iron as supported by higher maximum rates of inorganic iron uptake (Fig. 2.2d).

One exception to this increase in transporter expression was in centric diatoms at Experiment 1 where decreased abundances of ferric reductase proteins under increased CO₂ were observed (Fig. 2.4c). Ferric reductases have been implicated as part of a siderophore uptake pathway including that of Fe-DFB in diatoms (Coale et al. 2019), suggesting that this other pathway may not be favoured under these conditions; however, the ones expressed here may represent separate ferric reductase homologs that do not serve the same functional role. Ferrichrome-binding protein 1 (FBP1) acts as a receptor in concert with ferric reductases for siderophore uptake including Fe-DFB, but it was not differentially expressed (Supplemental Fig. 2.21 and 2.22) corresponding to the similar Fe-DFB uptake rates among treatments (Fig. 2.2d and Supplemental Fig. 2.6).

As some of the ferrous iron transporters with increased expression may be responsible for intracellular transport, these increases may also suggest a greater reliance on iron recycling or releasing stored iron to support growth in the short-term. Also in response to increased CO₂ at Experiment 1, centric diatoms increased transcript expression of a CREG-like gene (cellular repressor of E1-A stimulated genes; Fig. 2.4c) that is upregulated under iron stress in diatoms (Allen et al. 2008; Behnke et al. 2023; Lommer et al. 2012). Diatom CREG proteins are proposed to have a regulatory role in positively inducing endocytosis of iron when iron availability is low suggesting an increased dependence on endocytosis of iron under increased CO₂ (Turnšek et al. 2021).

Other differentially expressed genes indicate an adoption of strategies that reduce cellular iron demand. Due to its versatile redox chemistry, iron plays an essential role in photosynthesis-

related proteins in diatoms, and reductions in photosynthetic proteins can be related to iron stress (Coale et al. 2022). In centric diatoms, higher CO₂ reduced protein abundances of the photosynthetic proteins cytochrome b6 (petB) at Experiment 1 as well as apocytochrome f (petA) and photosystem II D1 (psbA) at Experiment 3 (Fig. 2.4c and Supplemental Fig. 2.23). Similarly, these proteins had lower abundances due to increased acidification in two previously mentioned low iron experiments in the Gulf of Alaska (Hopkinson et al. 2010). In pennate diatoms, lower protein abundances for the cytochrome b6f complex protein, petC, and the ferredoxin-NADP⁺ reductase, petH, were also found at Experiment 1.

Iron is also critical for mitochondrial electron transport (Coale et al. 2022), and in centric diatoms at Experiment 1, downregulation of an iron-containing mitochondrial protein, mitoNEET (mNT), that may be involved in signalling was observed (Fig. 2.4c)(Paddock et al. 2007). MitoNEET sequences were found to be widely distributed in diatoms where they are related to those in other eukaryotic phytoplankton including green algae (Supplementary Fig. 24). In the model green algae, *Chlamydomonas reinhardtii*, mNT was downregulated under iron stress indicating that this iron-sparing response may be more widely distributed among phytoplankton (Urzica et al. 2012).

Cellular iron requirements can also be reduced by using iron-free functionally equivalent proteins. In Experiment 3, centric diatoms increased transcript abundances of plastocyanin (petE), a copper-containing functional equivalent for the iron-requiring photosynthetic electron transport chain protein cytochrome c₆ (Peers and Price 2006), and increased protein abundances of the Cu-Zn utilizing superoxide dismutase (SOD; Fig. 4C). Different SOD proteins in diatoms employ different metal cofactors including iron (Groussman et al. 2015), and the

overrepresentation of Cu-Zn SOD may signify a preference for a non-iron using SOD under high CO₂.

Additionally, nitrogen assimilation proteins are estimated to account for a significant portion of the cellular iron requirement in phytoplankton (Raven 1988), and reducing their transcript or protein abundances has also been associated with iron stress in phytoplankton (Cohen et al. 2017). Under higher CO₂, pennate diatoms decreased transcript expression of ammonium transporters (*AMT*) and the ferredoxin-dependent glutamate synthase (*Fd-GOGAT*) in Experiment 1 and nitrate transporters (*NRT*) in Experiment 3 (Fig. 2.4c and Supplemental Fig. 2.25). Similarly, centric diatoms decreased transcript expression of the ferredoxin-dependent nitrite reductase (*Fd-Nir*) in Experiment 3 (Supplemental Figs. 2.23 and 2.25).

At the first time point of Experiment 3, *Ostreococcus* decreased transcript expression for the entire assimilatory nitrate reduction pathway including the plastid-localized nitrite transporter (*FNT*) and the nitrate transporter accessory protein, *NAR2* (Fig. 2.4c). *NRT* and *Fd-Nir* also displayed significantly lower protein abundances while *NAR2* abundances were not significantly different. *Fd-GOGAT* transcript expression was also lower at the first time point ($P = 0.052$, Fig. 2.4c and Supplemental Fig. 2.25), and by the second time point, protein abundances were significantly lower (Supplemental Figs. 2.25 and 2.26). *Ostreococcus* also reduced transcript expression at 800 ppm for these nitrate assimilation genes but not to the same degree as at 1200 ppm further indicating that responses are function of CO₂ (Supplemental Fig. 2.25). Additionally, *Ostreococcus* responded by increasing transcript expression of a SLC9-family Na⁺/H⁺ exchanger at 1200 ppm, potentially to regulate intracellular pH under higher acidity (Fig. 2.4c)(Fuster and Alexander 2014).

At the second time point, *Ostreococcus* displayed further iron-related responses (Supplemental Fig. 2.26). Like the increased protein abundances of the Cu-Zn SOD for centric diatoms, *Ostreococcus* increased expression of the Ni-using SOD. *Ostreococcus* also had reduced protein abundances of the iron storage protein ferritin (FTN) where it may play a role in buffering intracellular iron, a function that may no longer be beneficial under Fe stress (Botebol et al. 2015).

In Experiment 1, centric diatoms displayed remodelling of their carbon uptake pathway with changes in carbonic anhydrase (CA) transcript and protein abundances (Fig. 2.4c). Carbonic anhydrases (CA) interconvert CO₂ and HCO₃⁻ as part of the CO₂-concentrating mechanism (CCM) in diatoms to overcome relatively low concentrations of ambient CO₂ (Matsuda et al. 2017). Diatoms possess multiple CAs belonging to evolutionarily distinct families with diverse localizations including extracellular CAs at the cell surface, although there is not consistent conservation between family and location (Matsuda et al. 2017).

With increased CO₂, transcript expression of α -CA and β -CA in centric diatoms at Experiment 1 significantly increased while expression of δ -CA and ι -CA significantly decreased (Fig. 2.4c and Supplemental Fig. 2.23). These genes were not found to be differentially abundant on the protein level. Conversely, CAs in pennate diatoms at Experiment 1 were not significantly different on the transcript level, but δ -CA proteins significantly decreased with increased CO₂ aligning with the lower expression observed in centric diatoms.

In centric diatoms, α -CA, β -CA, and γ -CA expression was dominated by *Chaetoceros* whereas other centric diatom genera were expressing δ -CA and ι -CA indicating that these responses may be more genus-specific (Fig. 2.4c and Supplemental Fig. 2.23). As zinc is a common cofactor in carbonic anhydrases (Morel et al. 2020) and ZIP-family transporters may

also transport zinc (Blaby-Haas and Merchant 2012), increased expression of *ZIP1*, which was also largely from *Chaetoceros*, may serve to support increased α -CA and β -CA abundances under higher CO₂. α -CA and β -CA are believed to be localized to different compartments in the chloroplast in model diatoms (Matsuda et al. 2017), and expression of α -CA is thought to be constitutive (Samukawa et al. 2014). γ -CAs and δ -CAs have shown a wide range of localizations (Matsuda et al. 2017; Samukawa et al. 2014) and ι -CA has been localized to the chloroplast in *T. pseudonana* (Jensen et al. 2019). Both δ -CA and ι -CA have been shown to be downregulated under high CO₂ as observed here, although for δ -CA, this pattern is not entirely consistent (Crawford et al. 2011; Hennon et al. 2015; Jensen et al. 2019; Samukawa et al. 2014).

SLC4-family bicarbonate transporters are also part of the CCM (Nakajima et al. 2013). Despite change in CA expression, SLC4 transporter transcripts were not differentially expressed in any experiment for centric or pennate diatoms (Supplemental Fig. 2.27). In the protein data, it was only detected at Experiment 1 in centric diatoms after 72 hours where it was also not differentially abundant. *SLC4* expression has previously been shown to decrease with higher CO₂ in the model diatom *T. pseudonana* (Hennon et al. 2015); however, no difference in bicarbonate uptake or certain *SLC4* transcripts within a more extensive pH range (7.5 to 8.2) was observed in the model diatom *P. tricornutum* suggesting a lack of response over the pH ranges used here (Nakajima et al. 2013).

Discussion

Ocean acidification is hypothesized to alter iron bioavailability (McQuaid et al. 2018; Shi et al. 2010), which is particularly important as large areas of the ocean including some upwelling regions experience frequent iron limitation (Bruland et al. 2001; Hutchins et al. 2002; Till et al.

2019). Here we examined upwelling-associated phytoplankton communities' responses while they were simultaneously exposed to iron stress and acidification over short time periods. The iron status and responses of these communities were assessed with a combination of experimental, biogeochemical, physiological, and 'omics approaches. Although large differences were not observed in the bulk community, molecular-level responses with greater taxonomic resolution of dominant groups were indicative of acidification-driven alterations to iron bioavailability and greater iron stress under higher CO₂. Within centric and pennate diatoms, these responses were variable, but included changes to myriad of genes related to iron uptake and cellular iron demand (Fig. 2.5). Several of the iron-related genes described here are relatively uncharacterized in diatoms and of diverse evolutionary origins highlighting diatoms' unique adaptations to grow under low iron availability.

In Experiment 1, both centric and pennate diatoms employed iron uptake pathways for siderophore-bound iron or ferrous iron that are more bioavailable under acidic conditions than ferric iron (Fig. 2.4c, 2.5) (McQuaid et al. 2018; Shi et al. 2010). Here, a preference for siderophore-bound iron may not have been reflected in the uptake rates of Fe-DFB (Fig. 2.2d) as Fe-DFB uptake has been characterized with a different suite of genes than the ones observed to be upregulated here (*FecCD* and *SLC49A*). Genes responsible for Fe-DFB uptake, *FRE* and *FBPI*, were also not differentially expressed except for decreased protein abundances in centric diatoms in Experiment 1, although the ferric reductases expressed here may represent separate homologs that does not serve the same functional role (Fig. 2.4c and Supplemental Figs. 2.21 and 2.22)(Coale et al. 2019). Furthermore, uptake of Fe-DFB may be less affected by acidification compared to other siderophore complexes (McQuaid et al. 2018; Shi et al. 2010), and organically-complexed iron in seawater is generally more bioavailable (Shaked et al. 2020).

In Experiment 3, the response was instead transcriptional and proteomic increases in iron permeases (FTR) in centric diatoms signifying an increase in free inorganic iron transporters as its availability is reduced (Fig. 2.4c). This increase in transporters is further supported by higher maximal uptake rates of inorganic iron that are also a further sign of increased iron stress (Fig. 2.2d).

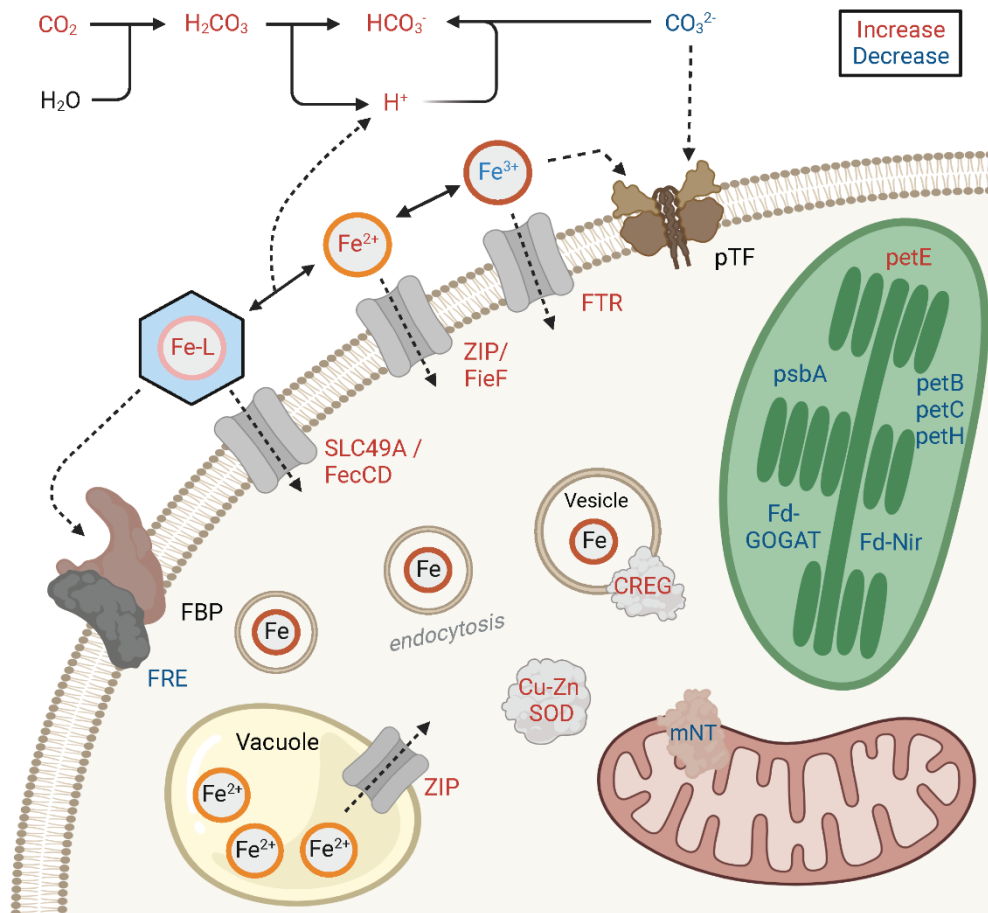


Figure 2.5: Potential diatom-related cellular changes from increased acidification driving changes in iron chemistry and/or bioavailability. Molecules or proteins are coloured red or blue for potential increases or decreases respectively. Black denotes no change. This figure was created with BioRender.com.

The difference in iron transporter expression between the two experiments may have occurred for several reasons. First, the iron status of the communities was distinctly different; the community at Experiment 1 was not iron-limited whereas the community at Experiment 3 was, although in Experiment 1, the $\text{NO}_3:\text{dFe}$ ratio suggested the potential for iron limitation and high expression of ISIP genes indicates some degree of iron stress (Fig. 2.4b). Meanwhile, Experiment 1 also had the highest initial dFe concentrations (Fig. 2.1b). Collectively, this may have allowed the community at Experiment 1 to possess relatively high intracellular and/or stored iron to rely on while subjected to added iron stress from acidification. The observed increased expression of ferrous iron transporters that are potentially intracellular may support this increased reliance on intracellular recycling or release of stored iron to sustain growth over these relatively short timescales (Blaby-Haas and Merchant 2012; Lampe et al. 2018). Second, the iron pools between the experiments were likely distinct although they are largely uncharacterized here. Besides higher dissolved iron concentrations, labile particulate iron concentrations were relatively high at Experiment 1, and dissolution of particulate iron may have supported growth (Fig. 2.1c and Supplemental Fig. 2.3). Third, the communities were distinct not only in terms of composition but also functional diversity (Fig. 2.3c). Only 12.2% and 21.5% of orthologous gene clusters in centric and pennate diatoms respectively were found to be shared between the two experiments (Supplemental Fig. 2.13). Certainly, there is high functional diversity within diatoms, and their responses may not be equivalent.

In both Experiments 1 and 3 however, diatoms also appeared to reduce their cellular iron requirements (Figs. 2.4c and 2.5, Supplemental Figs. 2.23 and 2.25). Decreased transcript or protein abundances of several iron-requiring proteins related to photosynthesis and nitrogen assimilation as well as a mitochondrial protein, mitoNEET, were observed. In Experiment 3,

centric diatoms increased transcript abundances of the iron-free protein plastocyanin and increased protein abundances of the Cu-Zn SOD which may substitute for the iron-using SOD.

Collectively, these changes in transcript and protein abundances signify an increase in iron stress resulting from acidification. The use of alternative iron uptake pathways or iron-sparing measures may have temporarily satisfied high cellular iron demand over the relatively short periods examined here, but longer-term exposure to high CO₂ may have resulted in a more iron-stressed community. For the community at Experiment 1, siderophore-bound iron or ferrous iron uptake pathways are likely less efficient than those for inorganic ferric iron uptake, and a reliance on iron storage is not sustainable over longer time scales. For the community at Experiment 3, the increase of iron permeases for ferric iron uptake confers increased energetic costs, and reductions in iron-requiring photosynthetic and nitrogen assimilation proteins may result in less efficiency in these critical pathways for growth. However, future iron supplies will likely be altered, and the net result remains unknown (Boyd et al. 2014; Capone and Hutchins 2013); increased iron inputs could offset reductions in bioavailability and prevent increased iron stress.

Remodeling of carbon uptake pathways was also observed in centric diatoms in Experiment 1 (Fig. 2.4c and Supplemental Fig. 2.23). Phytoplankton employ a carbon concentrating mechanism (CCM) that includes carbonic anhydrases (CA) to overcome relatively low concentrations of ambient CO₂ (Matsuda et al. 2017). Extracellular CAs (eCA) can alter the carbonate chemistry in the region immediately around the cell (Chrachri et al. 2018) with activities that span orders of magnitude and scale with cell size (Shen and Hopkinson 2015). This function may potentially overcome impacts from acidification such as maintaining elevated [CO₃²⁻] for iron uptake via phytoferritin. In Experiment 1, certain centric diatoms decreased

expression of CAs associated with low CO₂ as CO₂ increased which potentially translates into energetic savings (Mackey et al. 2015); however, *Chaetoceros* instead appeared to increase expression of α -CA and β -CA which would have the opposite effect (Fig. 2.4c). Potential energetic savings or costs may also be offset by other observed changes such as the relative increases in ribosomal proteins in pennate diatoms at Experiment 1 (Supplemental Fig. 2.16), other unknown proteins, or are negligible relative to the cell's energy demand contributing to the lack of large changes observed here.

The absence of large observable effects in Experiments 2 and 4 may be a result of reduced acidification exposure relative to community growth rates compared to Experiments 1 and 3 where multiple cell divisions may have occurred (Supplemental Fig. 2.8). Corresponding to this reduced exposure relative to growth rate, there were comparatively fewer transcriptional changes in Experiments 2 and 4 (Fig. 2.4a). Alternatively, there may have been an increase or no net change in bioavailable iron during those experiments. Under acidic conditions, the dissociation of iron from certain organic ligands is less favourable; however, this is not universal and some ligands are unaffected (Shi et al. 2010). Under more acidic conditions, some models indicate that Fe(III) may instead be more soluble leading to increases in bioavailable dissolved iron (Gledhill et al. 2015; Stockdale et al. 2016). Indeed, the iron pool and dynamics in the natural environment are vastly more complex than those in laboratory studies with both a single organism and iron source where reductions in iron bioavailability have been shown (Boyd et al. 2015).

Within all four experiments however, short-term exposure to pCO₂ levels surpassing global surface seawater means predicted by the end of the century and most natural surface water variability in the CCS (Supplemental Tables 2.1-2.3) produced few observable differences in

chlorophyll *a* production and nutrient uptake including that of both inorganic and organically-complexed iron (Fig. 2.2). Furthermore, increased CO₂ did not alter the carbon-to-nitrogen ratios of the community suggesting a lack of CO₂ fertilization with no increase in carbon content from high CO₂ as previously proposed (Fig. 2.2c) (Riebesell et al. 2007). Increased acidification did not create shifts in the prokaryotic or eukaryotic microbial communities over the short duration of the experiments as there were almost no significant differences in alpha diversity or community among treatments (Fig. 2.3b, 2.3c). Few ASVs (< 1.5%) were estimated to be differentially abundant (Fig. 2.3d), and of those that were, some may be calcifying phytoplankton that were impacted as calcification became less energetically favourable (Mackey et al. 2015).

These results may suggest that the natural phytoplankton communities observed here possess short-term resistance to acidification at pH values as low as 7.6 despite iron stress. Certainly, many phytoplankton taxa, especially diatoms, regularly bloom within the naturally acidic conditions of initially upwelled waters, and this resistance may have been selected for. Our experiment further offshore, Experiment 4, indicates that this lack of impact may extend farther than the nearshore region; however, these responses may still differ from regions where both pH variability and iron supply are even lower, e.g., high-nitrate low-chlorophyll regions. Again however, these results may have occurred from the relatively short duration of the experiments, and longer exposure may have showed either impacts from added iron stress or continued resistance.

Ocean acidification is in itself a multiple stressor as several carbon and iron parameters in seawater are changing simultaneously (Hurd et al. 2020). As a disturbance, it may be representative of a nonlinear transition where the tipping points are not reached until extreme

conditions are encountered (Scheffer et al. 2009). In the case of phytoplankton in upwelling regions, the tipping point may exist at atmospheric CO₂ levels beyond those estimated by the end of the century under high emissions scenarios or more prolonged exposure compared to the short time-scales used here (Supplementary Table 2). These extremes and interactive effects with additional stressors such as temperature for upwelling-associated phytoplankton communities remain to be uncovered.

Methods

Experimental design

Trace metal clean incubation experiments were conducted at four sites (Table 2.1) in August 2019 onboard the *R/V Atlantis* as part of the California Current Ecosystem Long-Term Ecological Research (CCE-LTER) Process Cruise. At each site at approximately 06:00 PST, seawater was collected from near the depth of maximum fluorescence (Supplemental Fig. 2.1) with a trace metal CTD-rosette fitted with Teflon coated 5 L Niskin-X bottles (Ocean Test Equipment) on a non-metallic hydroline. Niskin bottles were then brought into a positive pressure trace metal clean van where seawater was dispensed into 4 L polycarbonate bottles fitted with perfluoroalkoxy alkane (PFA) tubing and Kynar® PVDF barb fittings for sampling and bubbling. At the end of one inlet line, the bottles also had a plastic aerator (Lee's Discard-A-Stone) to maximize gas transfer.

Prior to the incubations, bottles were cleaned using trace metal techniques in a trace metal clean area under positive pressure. First, bottles were treated with acidic detergent (CitraNOX®, Alconox Inc.) for one day then rinsed five times with Milli-Q. Bottles were then filled with 1.2 mol L⁻¹ trace metal grade HCl for two weeks then rinsed again three times with

Milli-Q. During each step, detergent, Milli-Q, or HCl was syringed through the aerator, barbs, and tubing with all plastic syringes (NORM-JECT®).

Bottles were rinsed with seawater three times, then 4 L of seawater was dispensed into each bottle. To attempt to avoid macronutrient depletion, macronutrients (nitrate, phosphate, and silicic acid) that had been prepared with trace metal grade salts and passed over Chelex 100 resin to remove contaminating metals as described by Sunda et al. (2005). Macronutrients were added at varying concentrations based on estimated initial concentrations (Table 2.1). Bottles were then transferred to the laboratory onboard where they were placed in an incubator (Percival Scientific) at 12.5°C with 115 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ photosynthetically active radiation (PAR). The incubator was programmed on a light:dark cycle to closely match the timing of sunset and sunrise at the time and location of the incubations (lights on from 06:30 to 19:45). Seawater was then continuously bubbled with commercially prepared gas mixtures of air balanced with 400, 800, or 1200 ppm CO₂ (Praxair, Inc.) at a total rate of 0.3 LPM for each treatment. Gas was delivered to each bottle through a 4-port cross connector and fine-tuned with Keck™ tubing clamps to achieve evenly distributed flow of approximately 0.1 LPM in each bottle. Prior to the cross connector, gas was filtered through a 0.2 μm Millex-FG PTFE filter (Millipore), passed through tubing that also cleaned with 1.2 mol L⁻¹ trace metal clean HCl and rinsed three times with Milli-Q, then delivered to the inlet tubing at the bottle through another Whatman™ HEPA-Vent filter. A HEPA-Vent filter was also placed on the outlet tubing to prevent contaminating particles from entering the bottles. Triplicate bottles were used for all experimental conditions; one replicate from the 800 ppm treatment for experiment 4 (800B) was not included in final analyses due to potential iron contamination.

Seawater was collected directly from the Niskin bottles and immediately processed for the initial time point. For the later time points, sampling occurred at one-day intervals that varied among experiments: 48 and 72 hours for Experiment 1, 24 and 48 hours for Experiment 2, and 48 and 96 hours for Experiments 3 and 4 (Table 2.1). These intervals were chosen to attempt to avoid macronutrient depletion with maximum experimental duration of 96 hours to also avoid increased bottle effects. To sample these later timepoints or chlorophyll *a* at 24-hour intervals, bottles were brought into a HEPA-filtered positive pressure trace metal clean area. Samples were then immediately taken for dissolved inorganic carbon (DIC) followed by subsamples for chlorophyll *a*, particulate organic carbon and nitrogen, RNA, proteins, Fe uptake, macronutrient concentrations, and total alkalinity (TA) that were dispensed into additional bottles for sampling as described below.

Satellite data and maps

Satellite-derived data were obtained from the NOAA CoastWatch Browser, and maps were created with matplotlib v2.2.2 for Python v2.7 (Hunter 2007). Chlorophyll *a* concentrations on a 0.0125° grid were obtained from MODIS on board the Aqua (EOS PM).

Carbonate chemistry

Initial (T0) seawater samples for dissolved inorganic carbon (DIC) and total alkalinity (TA) analysis were collected from the Niskin bottles. Tubing was attached to the Niskin bottle and placed at the bottom of a 250 mL borosilicate glass bottle. The bottle was then inverted to thoroughly rinse the inside of the bottle and then filled from the bottom and allowed to overflow at least one full volume. The tubing was then pinched and withdrawn leaving only a small

headspace. 100 μL of a saturated mercuric chloride (HgCl_2) solution was added to the sample to prevent biological activity (Hunter 2007). Bottles were then sealed with high vacuum grease (Apiezon) on the stopper and fastened using a rubber band and bottle clip.

For the later timepoints, DIC and TA samples were collected separately. DIC was collected immediately by slowly syringing without introducing bubbles from the incubation bottles into 60- or 120-mL borosilicate glass bottles. Each bottle was gently rinsed with seawater then filled from the bottom using tubing connected to the syringe. Seawater was immediately poisoned with either 25 μL or 50 μL of saturated HgCl_2 solution for the 60- or 120-mL bottles respectively. Bottles were then sealed with grease and clipped as done for the T0 samples. DIC was measured with an Automated Infra-Red Inorganic Carbon Analyzer (AIRICA, Marianda) equipped with a LI-COR 7000 $\text{CO}_2/\text{H}_2\text{O}$ gas analyser as the detector.

TA was collected from filtrate from the sterivex used for RNA or proteins into 250 mL borosilicate glass bottles. Each bottle was rinsed thoroughly three times before filling. Seawater was poisoned with 100 μL saturated HgCl_2 solution and the bottles were sealed as described for the T0 samples. TA was then measured with an open-cell potentiometric acid titration (0.1 N HCl) system developed by the Dickson Lab at Scripps Institution of Oceanography (Dickson et al. 2007).

Both seawater DIC and TA measurements were made relative to certified reference materials (CRMs) provided by the laboratory of A. Dickson at Scripps Institution of Oceanography. The accuracy and precision of the measurements from the CRMs were $-0.21 \pm 1.26 \mu\text{mol kg}^{-1}$ ($n = 24$) for DIC and $-0.42 \pm 1.85 \mu\text{mol kg}^{-1}$ ($n = 17$) for TA. The complete seawater carbonic-acid system (i.e., pCO_2 , pH, CO_3^{2-} , HCO_3^- , etc) was calculated based on DIC, TA, temperature, salinity, phosphate, and silicate using CO2SYS for MATLAB (v 2.1)(Lewis

and Wallace 1998). Calculations used the first and second dissociation constants of carbonic acid (K_1 and K_2) from Mehrbach et al. (1973) refit by Dickson and Millero (1987), the dissociation constants of bisulfate (K_{HSO_4}) from Dickson (1990), and total boron from Uppstrom (1974). pH values are reported on the total H^+ scale.

Dissolved iron concentrations

Dissolved iron (dFe) was measured via chemiluminescence flow-injection analysis with hydrogen peroxide oxidation (FeLume, Waterville Analytical) as described by Lohan et al. (2006). Briefly, seawater was filtered using acid-cleaned 0.2 μm Acropak 200 capsule filters (VWR International) into clean and rinsed low-density polyethylene (Nalgene) bottles in a positive pressure clean van then was acidified to pH ~ 1.8 with HCl (Optima grade, Fisher Scientific). Dissolved iron was oxidized to Fe(III) with hydrogen peroxide (1% v/v optima grade), buffered in-line with an ammonium-acetate mixture, and selectively pre-concentrated on a column packed with resin (Toyopearl 650 M chelating resin) at pH ~ 3.5 . After the sample was eluted with 0.23 M HCl (Optima grade), the production of radicals catalyses a three-step oxidation reaction with hydrogen peroxide (0.23 M optima grade) and luminol (0.25 mM). Iron was then eluted and oxidized with pH > 9 luminol-ammonia buffer, and the chemiluminescence (425 nm) was measured with a photomultiplier tube. Final dissolved iron concentrations were then quantified via external standard curve. In-house and GEOTRACES consensus reference standards (geotraces.org) were measured in each analytical run to ensure accuracy and precision.

Labile particulate iron concentrations

Labile particulate iron was calculated as total dissolvable Fe (TDFe) minus dissolved iron (dFe). Unfiltered samples were acidified to pH $\sim$ 1.8 with HCl (Optima grade, Fisher Scientific) and stored for 22-24 months. Immediately prior to analysis, the samples were syringe-filtered using acid-cleaned syringes and 0.2 μ m Supor® filters, and then analyzed using the same analytical methods as for dFe samples. The blank from the syringe-filtration was averaged 0.058 nmol L⁻¹, which is at most a 6% correction to the overall measurement of TDFe. Due to the calibration range of the standards, it was necessary to dilute many of the samples, especially nearshore; therefore, offshore filtered Pacific surface seawater (0.22 nmol L⁻¹) was used to dilute samples by a factor of 10 prior to analysis.

Dissolved macronutrient concentrations

Approximately 30-35 mL of filtrate from the Sterivex filter unit (0.2 μ m) was dispensed into 50 mL polypropylene centrifuge tubes and frozen at -20°C until analysis. Each tube was rinsed three times before filling. Nutrients were analysed on a QuAATro continuous segmented flow autoanalyzer (SEAL Analytical) as described by the CalCOFI methods manual. Total oxidized nitrogen (nitrate reduced to nitrite), solely nitrite, and the efficiency of reduction of nitrate to nitrite were assessed to calculate solely the nitrate concentration. Reference materials for nutrients in seawater (KANSO Co., LTD.) were included in the run for quality control.

Si_{ex} proxy for diatom iron limitation

Si_{ex} is a biogeochemical proxy that suggests Fe limitation in diatoms due to preferential utilization of silicic acid (Hogle et al. 2018; Hutchins and Bruland 1998; King and Barbeau 2011), and was calculated as described by Hogle et al. (2018). Briefly,

$$Si_{ex} = [\mu mol\ H_4SiO_4\ L^{-1}] - ([\mu mol\ NO_3\ L^{-1}] * R_{Si:NO_3})$$

where $R_{Si:NO_3}$ is the preformed ratio of silicic acid and nitrate from upwelled source waters. Based on averaged values from the upwelling source water densities in the region ($\sigma_\theta = 25.8\ kg\ m^{-3}$), $R_{Si:NO_3} = 0.99$ (Hogle et al. 2018).

Parallel incubations to assess iron status

Parallel incubations were conducted to evaluate the presence of iron limitations. First, approximately four hours prior to the start of each of the pCO₂ experiments, seawater was collected from near the same depth as the pCO₂ experiments with the TMC rosette (Supplementary Fig. 1). One litre of seawater was dispensed into trace metal clean polycarbonate bottles. Trace metal clean techniques were used as described above. Triplicate bottles at each site were either unamended (control), supplemented with 5 nmol L⁻¹ iron (+Fe), or supplemented with 100 nmol L⁻¹ of the iron chelator desferrioxamine B (+DFB).

Bottles were incubated *in situ* for 24 hours with a quasi-Lagrangian drifter as described in Stukel et al. (2013). The drifter was equipped with a surface float and holey-sock drogue centred at 15 m. Bottles were attached to the drifter near the depth of collection in mesh bags. Either immediately (T0) or after 24 hours of incubation, samples were collected for chlorophyll *a* and macronutrients. The drifter was followed between the start of these incubations and the acidification incubations to sample from approximately the same water mass.

One day following the initiation of the fourth pCO₂ experiment (E4) and while still following the Lagrangian drifter, an additional trace metal clean incubation experiment was performed using a community collected and incubated on the drifter at 40 m depth. Triplicate 1 L polycarbonate bottles at each site were either unamended (control), supplemented with 5 nmol L⁻¹ (+Fe), supplemented with 10 μmol L⁻¹ (+N), or both iron and nitrate (+FeN). After 48 hours, samples were collected for chlorophyll *a*.

In addition to the incubations at E3 on the drifter, deckboard incubations in triplicate 2.7 L polycarbonate bottles were performed with neutral density screening for light and flow-through seawater for temperature. These bottles were either unamended (control) or supplemented with 5 nmol L⁻¹ (+Fe). Samples for chlorophyll *a* were collected immediately (T0) and following 24 and 48 hours of incubation.

Chlorophyll *a*

Fifty mL of seawater was filtered through a GF/B filter (1.0 μm, 25 mm) for the acidification experiments or a GF/F filter for the parallel incubations under gentle vacuum pressure (< 100 mm Hg) and frozen at -20°C until analysis. Chlorophyll *a* extraction was performed using 90% acetone at -20°C for 24 h and measured via *in vitro* fluorometry on a Turner Designs 10-AU fluorometer using the acidification method. Community growth rates (μ) were derived from the slope of linear regressions with the natural log of chlorophyll *a* concentrations over time. The last time point from Experiment 1 was excluded from the growth rate calculations as macronutrients were depleted.

Particulate organic carbon and nitrogen

Four hundred mL of seawater was filtered through a pre-combusted Whatman GF/B filter (1.0 μm , 25 mm, 450°C for 6 hours) under gentle vacuum pressure (< 100 mm Hg) and frozen in plastic petri dishes at -20. Filters were dried at 62°C then inorganic carbon was removed by acidifying the filters with HCl vapor followed by drying for another hour at 62°C. Filters were then encapsulated in tin and analysed on a Costech ECS 4010 CHNSO Analyzer following the manufacturer's instructions. Blank filters were run alongside the samples and subtracted from the sample quantities.

Short-term ^{59}Fe uptake assays

Iron uptake rates were characterized in subsamples collected from the incubation bottles or directly from the Niskin-X bottles (T0 samples) in acid-cleaned polycarbonate bottles. Additions of either 1 nmol L^{-1} $^{59}\text{FeCl}_3$ or 1 nmol L^{-1} ^{59}Fe -DFB (ferrioxamine B) were made to uptake bottles which were then incubated without light. ^{59}Fe -DFB stocks were equilibrated overnight with a 4:5 ratio of iron to DFB. Four aliquots of each uptake bottle were filtered onto 1 μm polycarbonate filters at one-hour intervals. Each filter was rinsed with an oxalate-EDTA wash solution for 15 minutes before rinsing with filtered seawater (Tang and Morel 2006). Filters were then preserved in Ecolite™ liquid scintillation cocktail (MP Biomedicals), and ^{59}Fe was quantified by liquid scintillation counting. Uptake rates were calculated by the linear regression of particulate ^{59}Fe per litre of seawater against time. Rates were normalized to both chlorophyll *a* and particulate organic carbon concentrations.

RNA collection and extraction

Between 525 and 1100 mL of seawater from each bottle or the initial seawater was filtered through a 0.22 µm Sterivex-GP filter unit (Cat. no. SVGP01050, Millipore) which was immediately sealed with a 1.2 mol L⁻¹ HCl and Milli-Q rinsed luer-lock plug and Hemato-SealTM tube sealant, wrapped in aluminum foil, and flash frozen in liquid nitrogen. RNA was extracted using the NucleoMag RNA extraction kit (Machery-Nagel) on an Eppendorf epMotion 5075t as described here: <https://dx.doi.org/10.17504/protocols.io.bd9ti96n>. RNA quality and quantity were assessed with an Agilent 2200 TapeStation (RNA ScreenTape) and the Quant-iTTM Ribogreen® RNA Assay kit. For RNA at initial time point (T0) for Experiment 3, *n* = 2.

16S and 18S rRNA Amplicon Sequencing and Analysis

From the aforementioned extracted RNA, cDNA was synthesized with the Invitrogen SuperScriptTM III First-Strand Synthesis System according to the manufacturer's instructions. RNA-derived amplicon libraries targeting the V4-V5 hypervariable region of the 16S gene and the V9 hypervariable region of the 18S region were then generated via a one-step PCR using the cDNA with the Azura TruFi DNA Polymerase PCR kit. For 16S, the 515-F (GTGYCAGCMGCCGCGGTAA) and 926R (CCGYCAATTYMTTTRAGTTT) primer set was used (Parada et al. 2016). For 18S, the 1389F (TTGTACACACCGCCC) and 1510R (CCTTCYGCAGGTTACCTAC) primer set was used (Amaral-Zettler et al. 2009). Each reaction was performed with an initial denaturing step at 95°C for 1 minute followed by 30 cycles of 95°C for 15 seconds, 56°C for 15 seconds, and 72°C for 30 seconds. 2.5 µL of each PCR reaction was ran on a 1.8% agarose gel confirm amplification. PCR products were purified using Beckman Coulter AMPure XP following the standard PCR clean-up protocol. PCR

quantification was performed using Invitrogen Quant-iT PicoGreen dsDNA Assay kit. Samples were then pooled in equal proportions (10 ng DNA per sample) followed by another AMPure XP PCR purification. The final libraries were evaluated on an Agilent 2200 TapeStation (D1000 ScreenTape), quantified with Qubit HS dsDNA kit, and each sequenced at the University of California, Davis Sequencing Core on a single Illumina MiSeq lane with a 15% PhiX spike-in.

Amplicons were analysed with QIIME2 v2019.10 (Bolyen et al. 2019). Briefly, demultiplexed paired-end reads were trimmed to remove adapter and primer sequences with cutadapt(Martin 2011). Trimmed reads were then denoised with DADA2 to produce amplicon sequence variants (ASVs)(Callahan et al. 2016). Taxonomic annotation of ASVs was conducted with the q2-feature-classifier classify-sklearn naïve-bayes classifier (Bokulich et al. 2018; Pedregosa et al. 2011) and the SILVA database (Release 138) for 16S (Pruesse et al. 2007) or the PR² database (v4.12.0) for 18S (Guillou et al. 2012). 16S ASVs with eukaryotic, plastid, and mitochondrial taxonomic assignments were removed. Diversity metrics and significance tests were calculated with the QIIME2 diversity plugin with samples rarefied to 3,756 reads for the 16S data and 82,249 reads for the 18S data (Supplementary Fig. 28)(Anderson 2001; Kruskal and Wallis 1952). Differential abundances of ASVs were assessed using the final time points of each experiment with DESeq2 v1.22.2 (Love et al. 2014; McMurdie and Holmes 2014; Weiss et al. 2017). ASVs were considered differentially abundant by those with Benjamini and Hochberg adjusted P-values < 0.05 (Benjamini and Hochberg 1995).

Metatranscriptome library preparation, sequencing, and analysis

A poly-A selected metatranscriptome sequencing library was constructed with the Illumina TruSeq Stranded mRNA kit according to the manufacturer's instructions and sequenced

on an Illumina NovaSeq 6000 with a S4 flow cell (2 x 150bp). Raw reads were quality trimmed with trimmomatic v0.39 (Bolger et al. 2014) and evaluated with FastQC v0.11.8 and MultiQC v1.9.dev0 (Ewels et al. 2016). Trimmed reads were assembled into contigs utilizing subsets of samples to generate sub-assemblies with Trinity v2.8.5 with *min_contig_length* set to 150 bases (Grabherr et al. 2011). These assemblies were then merged into final assemblies for each experiment with Trans-AbySS v2.0.1 (Robertson et al. 2010).

Proteins were predicted from the final assemblies with GenemarkS-T (Tang et al. 2015) which were deduplicated using the nucleotides of the predicted proteins and the dedupe script from BBMap v38.79. Ribosomal RNA sequences were identified with SortMeRNA v2.1 and were also removed (Kopylova et al. 2012). Reads were mapped to the predicted proteins with Bowtie2 v2.3.5 in local alignment mode (Langmead and Salzberg 2012). Only reads with a MAPQ score ≥ 8 were retained. Taxonomic annotation of predicted proteins was performed with DIAMOND (v0.9.31) BLASTP searches against PhyloDB v1.076 (Bertrand et al. 2015; Buchfink et al. 2015) with an E-value cutoff of 10^{-5} , minimum query coverage of 60%, and alignment scores in the top 95%. To avoid mis-annotations from contamination in the reference database, final taxonomic assignment was based on the Lineage Probability Index from the BLASTP hits with the aforementioned cutoffs (Bertrand et al. 2015; Podell and Gaasterland 2007). Orthogroups across different experiments were identified using the hierarchical phylogenetic approach implemented in OrthoFinder v2.5.4 (Emms and Kelly 2019). For functional annotation, DIAMOND BLASTP searches were performed against the Kyoto Encyclopedia of Genes and Genomes (KEGG; Release 94.1) (Bertrand et al. 2015; Kanehisa et al. 2017). KEGG Ortholog (KO) assignment was performed with KofamKOALA which utilizes hmmsearch against KOfam, an HMM database of KOs (Aramaki et al. 2019). Within taxonomic

groups, read counts were summed by KO annotation, and if KO annotation was absent, the KEGG gene annotation was used. Normalization and differential expression of transcripts were then assessed with DESeq2 v1.22.2. Significance in differentially expressed transcripts was determined with Benjamini and Hochberg adjusted *P*-values ($P < 0.05$) (Benjamini and Hochberg 1995). MA plots were generated with the code provided in Cohen et al. (2022).

Protein sample collection, extraction, and preparation

Seawater (300-900 mL) was filtered through a 0.22 µm Sterivex-GP filter unit (Cat. no. SVGP01050, Millipore) which was immediately sealed with a 1.2 mol L⁻¹ HCl and Milli-Q rinsed luer-lock plug and Hemato-SealTM tube sealant, wrapped in aluminium foil, and flash frozen in liquid nitrogen. Proteins were extracted with a detergent-based procedure based on Saito et al. (2011). First, filters were removed from the Sterivex casing as described by Cruaud et al. (2017) and placed in 750 µL of extraction buffer (2% SDS, 0.1 M Tris/HCl pH 7.5, 5 mM EDTA). Samples were incubated for 10 minutes at 95°C while mixed at 350 rpm. Sonication was then performed with a Diagenode Biorupter for approximately 8 minutes (15 second on/off intervals) at 4°C. Next, samples were again incubated at room temperature for 30 minutes while gently vortexing every 10 minutes. The filters were then removed from the tubes, and the remaining liquid was centrifuged at 15,000 x g for 30 minutes at room temperature. The supernatant containing the extracted protein was then removed without disturbing the pellet. Crude protein extracts were quantified with the PierceTM colorimetric BCA protein concentration assay.

S-TrapTM mini columns (PROTIFI) were then used to remove SDS extraction buffer from the samples and to digest extracted proteins into mass-spectrometry-compatible peptides.

Proteins were first reduced and alkylated using 5 mmol L⁻¹ dithiothreitol and 15 mmol L⁻¹ iodoacetamide respectively, and then denatured using 12% phosphoric acid. Samples were then diluted (1:7) with S-trapTM buffer (pH 7.1 100 mmol L⁻¹ tetraethylammonium bicarbonate in 90% aqueous methanol) and loaded onto the S-trapTM columns. To remove SDS, each sample was washed 10 times with 600 µL S-trapTM buffer using gentle vacuum. Proteins were then digested on-column using trypsin (1 trypsin: 25 protein) at 37°C for 16 hours. Tryptic peptides were eluted from the S-TrapTM columns with a series of 50 mmol L⁻¹ ammonium bicarbonate, 0.2 % aqueous formic acid, and 50% acetonitrile + 0.2% formic acid washes. Eluted samples were desalted on pre-conditioned 50 mg C18 columns (HyperSep), dried using a Vacufuge Plus (Eppendorf) and then resuspended in a 1% formic acid, 3% acetonitrile solution prior to liquid chromatography – tandem mass spectrometry analysis.

Metaproteomic processing and analysis

Samples were analyzed on a Dionex UltiMateTM 3000 RSLC-nano LC system (Thermo Scientific) coupled to a Q-Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific). Each sample was analyzed in duplicate, and a blank was run between each set of biological replicates. The peptide matrix was separated using one-dimensional liquid chromatography for 125 minutes under a non-linear gradient (Supplemental Table 2.6). The mass spectrometer was operated under positive polarity, data-dependent acquisition mode. Detailed mass spectrometer settings are shown in (Supplemental Table 2.7).

Raw mass spectrometry files were converted to mzML format using ThermoRawFileParser (Hulstaert et al. 2019). Two annotated databases were compiled using experiment-specific metatranscriptome sequences (see *Metatranscriptome library preparation*,

sequencing, and analysis) and appended with proteins from the common Repository of Adventitious Proteins to detect contaminants. Databases were then filtered to remove redundant sequences, and database searching was conducted using OpenMS MSGF+ with a 1% false discovery rate (Kim and Pevzner 2014; Röst et al. 2016). Post translational modification parameters included cysteine carbamidomethylation as a fixed modification, and methionine oxidation and N-terminal glutamate to pyroglutamate conversion as variable modifications.

Peptides were quantified using MS1 ion intensities, and FeatureFinderIdentification was used to identify unknown features based on identified MS2 spectra in replicate injections (Weisser and Choudhary 2017). Individual peptide abundances in each injection were normalized to the sum of all peptide intensities in that injection (total sum scaling), excluding peptides matching to the common contaminants database. Unique peptides were assigned to a taxonomic and functional annotation based on the metatranscriptome assembly and annotation. Non-unique peptides, i.e., peptides matching to more than one taxonomic group or functional annotation, were considered to be ambiguous. As with the transcripts, peptides were grouped based on KEGG Ortholog (KO) and gene annotations within each taxonomic group. Significant differences in protein group abundance for each taxonomic group were assessed within experiments using ion intensities with DESeq2 and Benjamini and Hochberg adjusted P-values ($P < 0.05$) (Benjamini and Hochberg 1995; Love et al. 2014). Although DESeq2 is designed for RNA-seq data, here it was used for consistency between the transcript and protein analyses, and it has been found to be effective at determining significance in differentially expressed proteins (Langley and Mayr 2015).

Ribosomal and photosynthetic protein mass fractions were calculated as described by McCain et al. (2022). Specifically, KO gene definitions were searched against character strings

for each function. The photosynthesis group included all proteins with “photosyn*”, “light-harvesting complex”, “flavodoxin”, or “plastocyanin” in their annotation where “*” represents a wildcard character. The ribosome group included all proteins with “ribosom*” and excluded proteins responsible for ribosomal synthesis. To be included, each taxonomic group of interest was required to have more than 50 unique peptides for each protein group to avoid known biases (McCain et al. 2022). Mass fractions were then calculated by summing the peptide abundances for each function within each taxonomic group and then dividing by the total peptide abundance for the taxonomic group. Differences in proteomic mass fractions were evaluated via ANOVAs followed by Tukey’s honest significant difference test in R.

Phylogenetic analyses

Taxonomically representative sets of homologs of diatom FTR, CDF-family, FecCD-domain, SLC49-family, and mitoNEET genes with an emphasis on unicellular eukaryotic algae were identified from the NCBI nr and MMETSP databases via BLASTP searches ($E\text{-value} \leq 10^{-3}$) (Keeling et al. 2014; Pruitt et al. 2007). To reduce the complexity stemming from the transcriptomes in MMETSP, sequences were clustered at the 95% similarity level with USEARCH (Edgar 2010). Seed sequences were then aligned using MAFFT v7 (Katoh and Standley 2013). Poorly aligned regions and regions rich in gaps were identified and removed using trimAI in *gappyout* mode (Capella-Gutiérrez et al. 2009). Alignments were then visualized and inspected in SeaView 4 (Gouy et al. 2010). Maximum likelihood phylogeny was computed in IQTree 2 under the model that produced the best fit using ModelFinder (Kalyaanamoorthy et al. 2017; Minh et al. 2020). Ultra-fast bootstrap approximation was estimated from 10,000 replicates.

Statistical analyses

Two-way ANOVAs followed by Tukey's honest significant difference test were performed on the biological and chemical properties of the seawater using Graphpad PRISM v9.0.0.

Data availability

The sequence data for the metatranscriptome and rRNA amplicon libraries reported in this study have been deposited in the National Center for Biotechnology (NCBI) sequence read archive under the BioProject accession no. PRJNA787648. Metatranscriptome assemblies, read counts, and annotations are available at Zenodo: <https://doi.org/10.5281/zenodo.5758778>. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD038549 (Perez-Riverol et al. 2019). The SILVA database was obtained from the SILVA website, and the PR² database was obtained from GitHub. The PhyloDB database is available on the A.E. Allen Lab website.

Acknowledgements

We thank the captain and crew of the *R/V Atlantis* as well as the participants of the California Current Ecosystem Long-Term Ecological Research (CCE LTER) P1908 Cruise (AT42-15). In particular, we thank Mark Ohman, Mike Stukel, and Ralf Goericke for their contributions to the overall cruise operations and Max Fenton who assisted with sample collection. We also thank Jeff McQuaid for advice on the experimental design.

This work was supported by the National Science Foundation Grants OCE-1756884 (to A.E.A.), OCE-1637632 (to K.A.B.), OCE-1756860 (to K.A.B. and A.J.A.), OCE-2224726 (to

K.A.B. and A.E.A.), and DGE-1650112 (Graduate Research Fellowship to R.H.L.), the National Ocean and Atmospheric Administration Grant NA19NOS4780181 (to A.E.A.), the Simons Foundation Collaboration on Principles of Microbial Ecosystems (PriME) Grant 970820 (to A.E.A.), the Simons Foundation Grant 504183 (to E.M.B.), the National Sciences and Engineering Research Council of Canada Grant RGPIN-2015-05009 (to E.M.B.), the Czech Science Foundation Grant 21-03224S (to M.O.), and the European Regional Development Fund ERDF/ESF Grant CZ.02.1.01./0.0/0.0/16_019/0000759.R.H.L. was also partially supported by a Chateaubriand Fellowship from the Office of Science and Technology of the Embassy of France in the United States and Inserm. This work also used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant ACI-1548562. Specifically, the Comet cluster at the San Diego Supercomputer Center was used through allocation TG-OCE200007 (to R.H.L.).

Chapter 2, in full, is a reprint of the material as it appears in Lampe, R.H., Coale, T.H., Forsch, K.O., Jabre, L.J., Kekuwa S., Bertrand, E.M., Horák A., Oborník M., Rabines, A.J., Rowland, E., Zheng, H., Andersson, A.J., Barbeau, K.A., and Allen, A.E. 2023. *Nature Communications* 14: 7215. The dissertation author was the primary investigator and author of this paper.

Author Contributions

R.H.L., T.H.C., K.O.F., E.M.B., S.K., A.J.A., K.A.B., A.E.A. designed the research; R.H.L., T.H.C., K.O.F., L.J.J., S.K., A.H., M.O., E.R., A.J.R., H.Z., K.A.B. performed the research; R.H.L., T.H.C., K.O.F., L.J.J., S.K., E.M.B., A.H., M.O., K.A.B., A.E.A. analysed the

data; R.H.L. drafted the manuscript; all authors contributed to the discussion of results and preparation of the manuscript.

Competing Interests

The authors declare no competing interests.

References

- Allen, A. E., J. Laroche, U. Maheswari, M. Lommer, N. Schauer, P. J. Lopez, G. Finazzi, A. R. Fernie, and C. Bowler. 2008. Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation. *Proc Natl Acad Sci* **105**: 10438-10443.
- Amaral-Zettler, L. A., E. A. McCliment, H. W. Ducklow, and S. M. Huse. 2009. A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLOS ONE* **4**: e6372.
- Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26**: 32-46.
- Aramaki, T., R. Blanc-Mathieu, H. Endo, K. Ohkubo, M. Kanehisa, S. Goto, and H. Ogata. 2019. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**: 2251-2252.
- Avendaño, L., M. Gledhill, E. P. Achterberg, V. M. C. Rérolle, and C. Schlosser. 2016. Influence of Ocean Acidification on the Organic Complexation of Iron and Copper in Northwest European Shelf Seas; a Combined Observational and Model Study. *Front Mar Sci* **3**, 58.
- Bach, L. T., and J. Taucher. 2019. CO₂ effects on diatoms: a synthesis of more than a decade of ocean acidification experiments with natural communities. *Ocean Sci* **15**: 1159-1175.
- Behnke, J., Y. Cai, H. Gu, and J. LaRoche. 2023. Short-term response to iron resupply in an iron-limited open ocean diatom reveals rapid decay of iron-responsive transcripts. *PLoS ONE* **18**: e0280827.
- Behnke, J., and J. LaRoche. 2020. Iron uptake proteins in algae and the role of Iron Starvation-Induced Proteins (ISIPs). *Eur J Phycol* **55**: 339-360.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* **57**: 289-300.
- Bertrand, E. M., J. P. McCrow, A. Moustafa, H. Zheng, J. B. McQuaid, T. O. Delmont, A. F. Post, R. E. Sipler, J. L. Spackeen, K. Xu, D. A. Bronk, D. A. Hutchins, and A. E. Allen. 2015. Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc Natl Acad Sci* **112**: 9938-9943.
- Beszteri, S., S. Thoms, V. Benes, L. Harms, and S. Trimborn. 2018. The response of three Southern Ocean phytoplankton species to ocean acidification and light availability: a transcriptomic study. *Protist* **169**: 958-975.
- Blaby-Haas, C. E., and S. S. Merchant. 2012. The ins and outs of algal metal transport. *Biochim Biophys Acta Mol Cell Res* **1823**: 1531-1552.
- Blaby-Haas, C. E., and S. S. Merchant. 2019. Comparative and Functional Algal Genomics. *Annu Rev Plant Biol* **70**: 605-638.

Bokulich, N. A., B. D. Kaehler, J. R. Rideout, M. Dillon, E. Bolyen, R. Knight, G. A. Huttley, and J. Gregory Caporaso. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**: 90.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics* **30**: 2114-2120.

Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. Cope, R. Da Silva, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciulek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Priesse, L. B. Rasmussen, A. Rivers, I. I. M. S. Robeson, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, and J. G. Caporaso. 2019. Reproducible, interactive, scalable, and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852-857.

Borer, P. M., B. Sulzberger, P. Reichard, and S. M. Kraemer. 2005. Effect of siderophores on the light-induced dissolution of colloidal iron (III)(hydr) oxides. *Mar Chem* **93**: 179-193.

Botebol, H., E. Lesuisse, R. Šuták, C. Six, J.-C. Lozano, P. Schatt, V. Vergé, A. Kirilovsky, J. Morrissey, T. Léger, J.-M. Camadro, A. Gueneugues, C. Bowler, S. Blain, and F.-Y. Bouget. 2015. Central role for ferritin in the day/night regulation of iron homeostasis in marine phytoplankton. *Proc Natl Acad Sci* **112**: 14652-14657.

Boyd, P. W., S. T. Lennartz, D. M. Glover, and S. C. Doney. 2014. Biological ramifications of climate-change-mediated oceanic multi-stressors. *Nat Clim Change* **5**: 71.

Boyd, P. W., R. F. Strzepek, M. J. Ellwood, D. A. Hutchins, S. D. Nodder, B. S. Twining, and S. W. Wilhelm. 2015. Why are biotic iron pools uniform across high- and low-iron pelagic ecosystems? *Global Biogeochem Cy* **29**: 1028-1043.

Breitbarth, E., R. J. Bellerby, C. C. Neill, M. V. Ardelan, M. Meyerhöfer, E. Zöllner, P. L. Croot, and U. Riebesell. 2010. Ocean acidification affects iron speciation during a coastal seawater mesocosm experiment. *Biogeosciences* **7**: 1065-1073.

Bruland, K. W., E. L. Rue, and G. J. Smith. 2001. Iron and macronutrients in California coastal upwelling regimes: Implications for diatom blooms. *Limnol Oceanogr* **46**: 1661-1674.

- Buchan, A., J. M. González, and M. A. Moran. 2005. Overview of the Marine *Roseobacter* Lineage. *Applied and Environmental Microbiology* **71**: 5665-5677.
- Buchfink, B., C. Xie, and D. H. Huson. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**: 59-60.
- Bundy, R. M., D. V. Biller, K. N. Buck, K. W. Bruland, and K. A. Barbeau. 2014. Distinct pools of dissolved iron-binding ligands in the surface and benthic boundary layer of the California Current. *Limnol Oceanogr* **59**: 769-787.
- Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**: 581-583.
- Capella-Gutiérrez, S., J. M. Silla-Martínez, and T. Gabaldón. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972-1973.
- Capone, D. G., and D. A. Hutchins. 2013. Microbial biogeochemistry of coastal upwelling regimes in a changing ocean. *Nature Geosci* **6**: 711-717.
- Chen, M., W.-X. Wang, and L. Guo. 2004. Phase partitioning and solubility of iron in natural seawater controlled by dissolved organic matter. *Global Biogeochem Cy* **18**.
- Chrachri, A., B. M. Hopkinson, K. Flynn, C. Brownlee, and G. L. Wheeler. 2018. Dynamic changes in carbonate chemistry in the microenvironment around single marine phytoplankton cells. *Nat Commun* **9**: 74.
- Coale, T. H., E. M. Bertrand, R. H. Lampe, and A. E. Allen. 2022. Molecular Mechanisms Underlying Micronutrient Utilization in Marine Diatoms, p. 567-604. *In* A. Falciatore and T. Mock [eds.], *The Molecular Life of Diatoms*. Springer International Publishing.
- Coale, T. H., M. Moosburner, A. Horák, M. Oborník, K. A. Barbeau, and A. E. Allen. 2019. Reduction-dependent siderophore assimilation in a model pennate diatom. *Proc Natl Acad Sci* **116**: 23609-23617.
- Cohen, N. R., H. Alexander, A. I. Krinos, S. K. Hu, and R. H. Lampe. 2022. Marine Microeukaryote Metatranscriptomics: Sample Processing and Bioinformatic Workflow Recommendations for Ecological Applications. *Front Mari Sci* **9**, 867007.
- Cohen, N. R., K. A. Ellis, R. H. Lampe, H. M. McNair, B. S. Twining, M. A. Brzezinski, M. T. Maldonado, F. I. Kuzminov, K. Thamtrakoln, C. P. Till, K. W. Bruland, W. G. Sunda, S. Bargu, and A. Marchetti. 2017. Variations in diatom transcriptional responses to changes in iron availability across ocean provinces. *Front Mar Sci* **4**: 360.
- Crawford, K. J., J. A. Raven, G. L. Wheeler, E. J. Baxter, and I. Joint. 2011. The Response of *Thalassiosira pseudonana* to Long-Term Exposure to Increased CO₂ and Decreased pH. *PLOS ONE* **6**: e26695.

- Cruaud, P., A. Vigneron, M.-S. Fradette, S. J. Charette, M. J. Rodriguez, C. C. Dorea, and A. I. Culley. 2017. Open the Sterivex™ casing: An easy and effective way to improve DNA extraction yields. *Limnol Oceanogr Methods* **15**: 1015-1020.
- Demir-Hilton, E., S. Sudek, M. L. Cuvelier, C. L. Gentemann, J. P. Zehr, and A. Z. Worden. 2011. Global distribution patterns of distinct clades of the photosynthetic picoeukaryote *Ostreococcus*. *ISME J* **5**: 1095-1107.
- Dickson, A., and F. J. Millero. 1987. A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media. *Deep Sea Res Part A. Oceanogr Res Pap* **34**: 1733-1743.
- Dickson, A. G. 1990. Thermodynamics of the dissociation of boric acid in synthetic seawater from 273.15 to 318.15 K. *Deep Sea Res Part A. Oceanogr Res* **37**: 755-766.
- Dickson, A. G., C. L. Sabine, and J. R. Christian. 2007. Guide to best practices for ocean CO₂ measurements. North Pacific Marine Science Organization.
- Dupont, C. L., J. P. McCrow, R. Valas, A. Moustafa, N. Walworth, U. Goodenough, R. Roth, S. L. Hogle, J. Bai, Z. I. Johnson, E. Mann, B. Palenik, K. A. Barbeau, J. C. Venter, and A. E. Allen. 2015. Genomes and gene expression across light and productivity gradients in eastern subtropical Pacific microbial communities. *ISME J* **9**: 1076-1092.
- Dutkiewicz, S., J. J. Morris, M. J. Follows, J. Scott, O. Levitan, S. T. Dyhrman, and I. Berman-Frank. 2015. Impact of ocean acidification on the structure of future phytoplankton communities. *Nat Clim Change* **5**: 1002-1006.
- Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-2461.
- Edwardsen, B., W. Eikrem, J. Green, R. A. Andersen, S. Y. Moon-van der Staay, and L. K. Medlin. 2000. Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. *Phycologia* **39**: 19-35.
- Emms, D. M., and S. Kelly. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* **20**: 238.
- Endo, H., K. Sugie, T. Yoshimura, and K. Suzuki. 2015. Effects of CO₂ and iron availability on *rbcL* gene expression in Bering Sea diatoms. *Biogeosciences* **12**: 2247-2259.
- Endo, H., T. Yoshimura, T. Kataoka, and K. Suzuki. 2013. Effects of CO₂ and iron availability on phytoplankton and eubacterial community compositions in the northwest subarctic Pacific. *J Exp Mar Biol Ecol* **439**: 160-175.
- Ewels, P., M. Magnusson, S. Lundin, and M. Käller. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**: 3047-3048.

- Feely, R. A., C. L. Sabine, J. M. Hernandez-Ayon, D. Ianson, and B. Hales. 2008. Evidence for Upwelling of Corrosive "Acidified" Water onto the Continental Shelf. *Science* **320**: 1490-1492.
- Feng, Y., C. Hare, J. Rose, S. Handy, G. DiTullio, P. Lee, W. Smith Jr, J. Peloquin, S. Tozzi, and J. Sun. 2010. Interactive effects of iron, irradiance and CO₂ on Ross Sea phytoplankton. *Deep Sea Res Part I Oceanogr Res Pap* **57**: 368-383.
- Fiechter, J., E. N. Curchitser, C. A. Edwards, F. Chai, N. L. Goebel, and F. P. Chavez. 2014. Air-sea CO₂ fluxes in the California Current: Impacts of model resolution and coastal topography. *Global Biogeochem Cy* **28**: 371-385.
- Forsch, K. O., K. C. Fulton, M. M. Weiss, J. W. Krause, M. R. Stukel, and K. A. Barbeau. 2023. Iron limitation and biogeochemical effects in southern California Current coastal upwelling filaments. *J Geophys Res Oceans* **128**: e2023JC019961.
- Fuster, D. G., and R. T. Alexander. 2014. Traditional and emerging roles for the SLC9 Na⁺/H⁺ exchangers. *Pflügers Arch* **466**: 61-76.
- Gledhill, M., E. P. Achterberg, K. Li, K. N. Mohamed, and M. J. A. Rijkenberg. 2015. Influence of ocean acidification on the complexation of iron and copper by organic ligands in estuarine waters. *Mar Chem* **177**: 421-433.
- Gledhill, M., C. M. G. van den Berg, R. F. Nolting, and K. R. Timmermans. 1998. Variability in the speciation of iron in the northern North Sea. *Mar Chem* **59**: 283-300.
- Gouy, M., S. Guindon, and O. Gascuel. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221-224.
- Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech* **29**: 644-652.
- Groussman, R. D., M. S. Parker, and E. V. Armbrust. 2015. Diversity and Evolutionary History of Iron Metabolism Genes in Diatoms. *PLoS ONE* **10**: e0129081.
- Gruber, N., D. Clement, B. R. Carter, R. A. Feely, S. van Heuven, M. Hoppema, M. Ishii, R. M. Key, A. Kozyr, S. K. Lauvset, C. Lo Monaco, J. T. Mathis, A. Murata, A. Olsen, F. F. Perez, C. L. Sabine, T. Tanhua, and R. Wanninkhof. 2019. The oceanic sink for anthropogenic CO₂ from 1994 to 2007. *Science* **363**: 1193-1199.
- Guillou, L., D. Bachar, S. Audic, D. Bass, C. Berney, L. Bittner, C. Boutte, G. Burgaud, C. de Vargas, J. Decelle, J. del Campo, J. R. Dolan, M. Dunthorn, B. Edvardsen, M. Holzmann, W. H. C. F. Kooistra, E. Lara, N. Le Bescot, R. Logares, F. Mahé, R. Massana, M. Montresor, R. Morard, F. Not, J. Pawlowski, I. Probert, A.-L. Sauvadet, R. Siano, T. Stoeck, D. Vaultot, P. Zimmermann, and R. Christen. 2012. The Protist Ribosomal Reference database (PR2): a catalog

of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* **41**: D597-D604.

Harrison, G. I., and F. M. Morel. 1986. Response of the marine diatom *Thalassiosira weissflogii* to iron stress. *Limnol Oceanogr* **31**: 989-997.

Hauri, C., N. Gruber, M. Vogt, S. C. Doney, R. A. Feely, Z. Lachkar, A. Leinweber, A. M. P. McDonnell, M. Munnich, and G. K. Plattner. 2013. Spatiotemporal variability and long-term trends of ocean acidification in the California Current System. *Biogeosciences* **10**: 193-216.

Hennon, G. M. M., J. Ashworth, R. D. Groussman, C. Berthiaume, R. L. Morales, N. S. Baliga, M. V. Orellana, and E. V. Armbrust. 2015. Diatom acclimation to elevated CO₂ via cAMP signalling and coordinated gene expression. *Nat Clim Change* **5**: 761-765.

Hoarfrost, A., S. Nayfach, J. Ladau, S. Yooseph, C. Arnosti, C. L. Dupont, and K. S. Pollard. 2020. Global ecotypes in the ubiquitous marine clade SAR86. *ISME J* **14**: 178-188.

Hogle, S. L., C. L. Dupont, B. M. Hopkinson, A. L. King, K. N. Buck, K. L. Roe, R. K. Stuart, A. E. Allen, E. L. Mann, Z. I. Johnson, and K. A. Barbeau. 2018. Pervasive iron limitation at subsurface chlorophyll maxima of the California Current. *Proc Nat Acad Sci* **115**: 13300-13305.

Hopkinson, B. M., Y. Xu, D. Shi, P. J. McGinn, and F. M. M. Morel. 2010. The effect of CO₂ on the photosynthetic physiology of phytoplankton in the Gulf of Alaska. *Limnol Oceanogr* **55**: 2011-2024.

Hoppe, C. J. M., C. S. Hassler, C. D. Payne, P. D. Tortell, B. Rost, and S. Trimborn. 2013. Iron Limitation Modulates Ocean Acidification Effects on Southern Ocean Phytoplankton Communities. *PLOS ONE* **8**: e79890.

Hulstaert, N., J. Shofstahl, T. Sachsenberg, M. Walzer, H. Barsnes, L. Martens, and Y. Perez-Riverol. 2019. ThermoRawFileParser: modular, scalable, and cross-platform RAW file conversion. *J Proteome Res* **19**: 537-542.

Hunter, J. D. 2007. Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* **9**: 90-95.

Hurd, C. L., J. Beardall, S. Comeau, C. E. Cornwall, J. N. Havenhand, P. L. Munday, L. M. Parker, J. A. Raven, and C. M. McGraw. 2020. Ocean acidification as a multiple driver: how interactions between changing seawater carbonate parameters affect marine life. *Mar Freshw Res* **71**: 263-274.

Hutchins, D., C. Hare, R. Weaver, Y. Zhang, G. Firme, G. DiTullio, M. Alm, S. Riseman, J. Maucher, and M. Geesey. 2002. Phytoplankton iron limitation in the Humboldt Current and Peru Upwelling. *Limnol Oceanogr* **47**: 997-1011.

Hutchins, D. A., and P. W. Boyd. 2016. Marine phytoplankton and the changing ocean iron cycle. *Nat Clim Change* **6**: 1072.

- Hutchins, D. A., and K. W. Bruland. 1998. Iron-limited diatom growth and Si:N uptake ratios in a coastal upwelling regime. *Nature* **393**: 561-564.
- Hutchins, D. A., A. E. Witter, A. Butler, and G. W. Luther. 1999. Competition among marine phytoplankton for different chelated iron species. *Nature* **400**: 858-861.
- Ibuot, A., A. P. Dean, and J. K. Pittman. 2020. Multi-genomic analysis of the cation diffusion facilitator transporters from algae. *Metallomics* **12**: 617-630.
- James, C. C., A. D. Barton, L. Z. Allen, R. H. Lampe, A. Rabines, A. Schulberg, H. Zheng, R. Goericke, K. D. Goodwin, and A. E. Allen. 2022. Influence of nutrient supply on plankton microbiome biodiversity and distribution in a coastal upwelling region. *Nature Commun* **13**: 2448.
- Jensen, E. L., R. Clement, A. Kosta, S. C. Maberly, and B. Gontero. 2019. A new widespread subclass of carbonic anhydrase in marine phytoplankton. *The ISME Journal* **13**: 2094-2106.
- Joint, I., S. C. Doney, and D. M. Karl. 2011. Will ocean acidification affect marine microbes? *ISME J* **5**: 1-7.
- Jones, B. M., M. D. Iglesias-Rodriguez, P. J. Skipp, R. J. Edwards, M. J. Greaves, J. R. Young, H. Elderfield, and C. D. O'Connor. 2013. Responses of the *Emiliania huxleyi* Proteome to Ocean Acidification. *PLOS ONE* **8**: e61868.
- Joos, F., and R. Spahni. 2008. Rates of change in natural and anthropogenic radiative forcing over the past 20,000 years. *Proc Nat Acad Sci* **105**: 1425-1430.
- Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermin. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**: 587-589.
- Kanehisa, M., M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353-D361.
- Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772-780.
- Kazamia, E., R. Sutak, J. Paz-Yepes, R. G. Dorrell, F. R. J. Vieira, J. Mach, J. Morrissey, S. Leon, F. Lam, E. Pelletier, J.-M. Camadro, C. Bowler, and E. Lesuisse. 2018. Endocytosis-mediated siderophore uptake as a strategy for Fe acquisition in diatoms. *Sci Adv* **4**: eaar4536.
- Keeling, C. D., S. C. Piper, R. B. Bacastow, M. Wahlen, T. P. Whorf, M. Heimann, and H. A. Meijer. 2005. Atmospheric CO₂ and ¹³CO₂ exchange with the terrestrial biosphere and oceans from 1978 to 2000: Observations and carbon cycle implications, p. 83-113. *A history of atmospheric CO₂ and its effects on plants, animals, and ecosystems*. Springer.
- Keeling, P. J., F. Burki, H. M. Wilcox, B. Allam, E. E. Allen, L. A. Amaral-Zettler, E. V. Armbrust, J. M. Archibald, A. K. Bharti, C. J. Bell, B. Beszteri, K. D. Bidle, C. T. Cameron, L.

- Campbell, D. A. Caron, R. A. Cattolico, J. L. Collier, K. Coyne, S. K. Davy, P. Deschamps, S. T. Dyhrman, B. Edvardsen, R. D. Gates, C. J. Gobler, S. J. Greenwood, S. M. Guida, J. L. Jacobi, K. S. Jakobsen, E. R. James, B. Jenkins, U. John, M. D. Johnson, A. R. Juhl, A. Kamp, L. A. Katz, R. Kiene, A. Kudryavtsev, B. S. Leander, S. Lin, C. Lovejoy, D. Lynn, A. Marchetti, G. McManus, A. M. Nedelcu, S. Menden-Deuer, C. Miceli, T. Mock, M. Montresor, M. A. Moran, S. Murray, G. Nadathur, S. Nagai, P. B. Ngam, B. Palenik, J. Pawlowski, G. Petroni, G. Piganeau, M. C. Posewitz, K. Rengefors, G. Romano, M. E. Rumpho, T. Rynearson, K. B. Schilling, D. C. Schroeder, A. G. B. Simpson, C. H. Slamovits, D. R. Smith, G. J. Smith, S. R. Smith, H. M. Sosik, P. Stief, E. Theriot, S. N. Twary, P. E. Umale, D. Vaultot, B. Wawrik, G. L. Wheeler, W. H. Wilson, Y. Xu, A. Zingone, and A. Z. Worden. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLOS Biol* **12**: e1001889.
- Khan, A. A., and J. G. Quigley. 2013. Heme and FLVCR-related transporter families SLC48 and SLC49. *Mol Aspects Med* **34**: 669-682.
- Kim, S., and P. A. Pevzner. 2014. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat Commun* **5**: 1-10.
- King, A. L., and K. A. Barbeau. 2011. Dissolved iron and macronutrient distributions in the southern California Current System. *J Geophys Res Oceans* **116**.
- Kopylova, E., L. No  , and H. Touzet. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211-3217.
- Kruskal, W. H., and W. A. Wallis. 1952. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc* **47**: 583-621.
- Kustka, A. B., A. E. Allen, and F. M. M. Morel. 2007. Sequence analysis and transcriptional regulation of iron acquisition genes in two marine diatoms. *J Phycol* **43**: 715-729.
- Lampe, R. H., E. L. Mann, N. R. Cohen, C. P. Till, K. Thamatrakoln, M. A. Brzezinski, K. W. Bruland, B. S. Twining, and A. Marchetti. 2018. Different iron storage strategies among bloom-forming diatoms. *Proc Natl Acad Sci* **115**: E12275-E12284.
- Langley, S. R., and M. Mayr. 2015. Comparative analysis of statistical methods used for detecting differential expression in label-free mass spectrometry proteomics. *J Proteom* **129**: 83-92.
- Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**: 357-359.
- Leinweber, A., and N. Gruber. 2013. Variability and trends of ocean acidification in the Southern California Current System: A time series from Santa Monica Bay. *J Geophys Res Oceans* **118**: 3622-3633.
- Lewis, E., and D. Wallace. 1998. Program developed for CO₂ system calculations. Environmental System Science Data Infrastructure for a Virtual Ecosystem.

- Lohan, M. C., A. M. Aguilar-Islas, and K. W. Bruland. 2006. Direct determination of iron in acidified (pH 1.7) seawater samples by flow injection analysis with catalytic spectrophotometric detection: Application and intercomparison. *Limnol Oceanogr-Meth* **4**: 164-171.
- Lommer, M., M. Specht, A.-S. Roy, L. Kraemer, R. Andreson, M. A. Gutowska, J. Wolf, S. V. Bergner, M. B. Schilhabel, U. C. Klostermeier, R. G. Beiko, P. Rosenstiel, M. Hippler, and J. LaRoche. 2012. Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol* **13**: R66.
- Lorenzo, M. R., M. Segovia, J. T. Cullen, and M. T. Maldonado. 2020. Particulate trace metal dynamics in response to increased CO₂ and iron availability in a coastal mesocosm experiment. *Biogeosciences* **17**: 757-770.
- Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.
- Mackey, K. R. M., J. J. Morris, F. M. M. Morel, and S. A. Kranz. 2015. Response of Photosynthesis to Ocean Acidification. *Oceanography* **28**: 74-91.
- Maldonado, M. T., and N. M. Price. 1999. Utilization of iron bound to strong organic ligands by plankton communities in the subarctic Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* **46**: 2447-2473.
- Maldonado, M. T., and N. M. Price. 2001. Reduction and transport of organically bound iron by *Thalassiosira Oceanica* (Bacillariophyceae). *J Phycol* **37**: 298-310.
- Malviya, S., E. Scalco, S. Audic, F. Vincent, A. Veluchamy, J. Poulain, P. Wincker, D. Iudicone, C. de Vargas, L. Bittner, A. Zingone, and C. Bowler. 2016. Insights into global diatom distribution and diversity in the world's ocean. *Proc Natl Acad Sci* **113**: E1516-E1525.
- Marchetti, A., and M. T. Maldonado. 2016. Iron, p. 233-279. *In* M. A. Borowitzka, J. Beardall and J. A. Raven [eds.], *The Physiology of Microalgae*. Springer International Publishing.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*; Vol 17, No 1: Next Generation Sequencing Data Analysis.
- Matsuda, Y., B. M. Hopkinson, K. Nakajima, C. L. Dupont, and Y. Tsuji. 2017. Mechanisms of carbon dioxide acquisition and CO₂ sensing in marine diatoms: a gateway to carbon metabolism. *Philos Trans R Soc Lond B* **372**: 20160403.
- Mausz, M. A., M. Segovia, A. Larsen, S. A. Berger, J. K. Egge, and G. Pohnert. 2020. High CO₂ concentration and iron availability determine the metabolic inventory in an *Emiliania huxleyi*-dominated phytoplankton community. *Environ Microbiol* **22**: 3863-3882.
- McCain, J. S. P., A. E. Allen, and E. M. Bertrand. 2022. Proteomic traits vary across taxa in a coastal Antarctic phytoplankton bloom. *ISME J* **16**: 569-579.

- McMurdie, P. J., and S. Holmes. 2014. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLOS Comput Biol* **10**: e1003531.
- McQuaid, J. B., A. B. Kustka, M. Oborník, A. Horák, J. P. McCrow, B. J. Karas, H. Zheng, T. Kindeberg, A. J. Andersson, K. A. Barbeau, and A. E. Allen. 2018. Carbonate-sensitive phytoferritin controls high-affinity iron uptake in diatoms. *Nature* **555**: 534.
- Mehrbach, C., C. Culbertson, J. Hawley, and R. Pytkowicz. 1973. Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure 1. *Limnol Oceanogr* **18**: 897-907.
- Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. Von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* **37**: 1530-1534.
- Moore, C. M., M. M. Mills, K. R. Arrigo, I. Berman-Frank, L. Bopp, P. W. Boyd, E. D. Galbraith, R. J. Geider, C. Guieu, S. L. Jaccard, T. D. Jickells, J. La Roche, T. M. Lenton, N. M. Mahowald, E. Marañón, I. Marinov, J. K. Moore, T. Nakatsuka, A. Oschlies, M. A. Saito, T. F. Thingstad, A. Tsuda, and O. Ulloa. 2013. Processes and patterns of oceanic nutrient limitation. *Nature Geosci* **6**: 701-710.
- Morel, F. M. M., P. J. Lam, and M. A. Saito. 2020. Trace Metal Substitution in Marine Phytoplankton. *Annual Review of Earth and Planetary Sciences* **48**: 491-517.
- Nakajima, K., A. Tanaka, and Y. Matsuda. 2013. SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc Nat Acad Sci* **110**: 1767-1772.
- Osborne, E. B., R. C. Thunell, N. Gruber, R. A. Feely, and C. R. Benitez-Nelson. 2020. Decadal variability in twentieth-century ocean acidification in the California Current Ecosystem. *Nat Geosci* **13**: 43-49.
- Osma, N., L. Latorre-Melín, B. Jacob, P. Y. Contreras, P. von Dassow, and C. A. Vargas. 2020. Response of Phytoplankton Assemblages From Naturally Acidic Coastal Ecosystems to Elevated pCO₂. *Front Mar Sci* **7**, 323.
- Paddock, M. L., S. E. Wiley, H. L. Axelrod, A. E. Cohen, M. Roy, E. C. Abresch, D. Capraro, A. N. Murphy, R. Nechushtai, J. E. Dixon, and P. A. Jennings. 2007. MitoNEET is a uniquely folded 2Fe-2S outer mitochondrial membrane protein stabilized by pioglitazone. *Proc Natl Acad Sci* **104**: 14342-14347.
- Parada, A. E., D. M. Needham, and J. A. Fuhrman. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* **18**: 1403-1414.
- Pausch, F., F. Koch, C. Hassler, A. Bracher, K. Bischof, and S. Trimborn. 2022. Responses of a Natural Phytoplankton Community From the Drake Passage to Two Predicted Climate Change Scenarios. *Front Mar Sci* **9**: 759501.

- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg. 2011. Scikit-learn: Machine learning in Python. *J Mach Learn Res* **12**: 2825-2830.
- Peers, G., and N. M. Price. 2006. Copper-containing plastocyanin used for electron transport by an oceanic diatom. *Nature* **441**: 341-344.
- Perez-Riverol, Y., A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, and M. Eisenacher. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**: D442-D450.
- Podell, S., and T. Gaasterland. 2007. DarkHorse: a method for genome-wide prediction of horizontal gene transfer. *Genome Biol* **8**: R16.
- Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188-7196.
- Pruitt, K. D., T. Tatusova, and D. R. Maglott. 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* **35**: D61-D65.
- Raven, J. A. 1988. The iron and molybdenum use efficiencies of plant growth with different energy, carbon and nitrogen sources. *New Phytol* **109**: 279-287.
- Riebesell, U., K. G. Schulz, R. G. J. Bellerby, M. Botros, P. Fritsche, M. Meyerhöfer, C. Neill, G. Nondal, A. Oschlies, J. Wohlers, and E. Zöllner. 2007. Enhanced biological carbon consumption in a high CO₂ ocean. *Nature* **450**: 545-548.
- Robertson, G., J. Schein, R. Chiu, R. Corbett, M. Field, S. D. Jackman, K. Mungall, S. Lee, H. M. Okada, J. Q. Qian, M. Griffith, A. Raymond, N. Thiessen, T. Cezard, Y. S. Butterfield, R. Newsome, S. K. Chan, R. She, R. Varhol, B. Kamoh, A.-L. Prabhu, A. Tam, Y. Zhao, R. A. Moore, M. Hirst, M. A. Marra, S. J. M. Jones, P. A. Hoodless, and I. Birol. 2010. De novo assembly and analysis of RNA-seq data. *Nat Meth* **7**: 909-912.
- Röst, H. L., T. Sachsenberg, S. Aiche, C. Bielow, H. Weissner, F. Aicheler, S. Andreotti, H.-C. Ehrlich, P. Gutenbrunner, and E. Kenar. 2016. OpenMS: a flexible open-source software platform for mass spectrometry data analysis. *Nat Methods* **13**: 741-748.
- Saito, M. A., V. V. Bulygin, D. M. Moran, C. Taylor, and C. Scholin. 2011. Examination of microbial proteome preservation techniques applicable to autonomous environmental sample collection. *Front Microbiol* **2**: 215-215.
- Samukawa, M., C. Shen, B. M. Hopkinson, and Y. Matsuda. 2014. Localization of putative carbonic anhydrases in the marine diatom, *Thalassiosira pseudonana*. *Photosynth Res* **121**: 235-249.

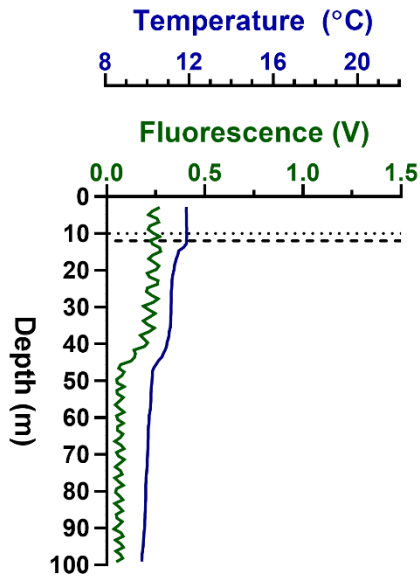
- Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara. 2009. Early-warning signals for critical transitions. *Nature* **461**: 53-59.
- Segovia, M., M. R. Lorenzo, M. T. Maldonado, A. Larsen, S. A. Berger, T. M. Tsagaraki, F. J. Lázaro, C. Iñiguez, C. García-Gómez, A. Palma, M. A. Mausz, F. J. L. Gordillo, J. A. Fernández, J. L. Ray, and J. K. Egge. 2017. Iron availability modulates the effects of future CO₂ levels within the marine planktonic food web. *Mar Ecol Prog Ser* **565**: 17-33.
- Shaked, Y., K. N. Buck, T. Mellett, and M. T. Maldonado. 2020. Insights into the bioavailability of oceanic dissolved Fe from phytoplankton uptake kinetics. *ISME J* **14**: 1182-1193.
- Shen, C., and B. M. Hopkinson. 2015. Size scaling of extracellular carbonic anhydrase activity in centric marine diatoms. *Journal of Phycology* **51**: 255-263.
- Shi, D., Y. Xu, B. M. Hopkinson, and F. M. M. Morel. 2010. Effect of Ocean Acidification on Iron Availability to Marine Phytoplankton. *Science* **327**: 676-679.
- Smith, S. R., J. T. F. Gillard, A. B. Kustka, J. P. McCrow, J. H. Badger, H. Zheng, A. M. New, C. L. Dupont, T. Obata, A. R. Fernie, and A. E. Allen. 2016. Transcriptional Orchestration of the Global Cellular Response of a Model Pennate Diatom to Diel Light Cycling under Iron Limitation. *PLoS Genet* **12**: e1006490.
- Staudenmaier, H., B. Van Hove, Z. Yaraghi, and V. Braun. 1989. Nucleotide sequences of the fecBCDE genes and locations of the proteins suggest a periplasmic-binding-protein-dependent transport mechanism for iron(III) dicitrate in *Escherichia coli*. *J Bacteriol* **171**: 2626.
- Stockdale, A., E. Tipping, S. Lofts, and R. J. G. Mortimer. 2016. Effect of Ocean Acidification on Organic and Inorganic Speciation of Trace Metals. *Environ Sci Technol* **50**: 1906-1913.
- Stukel, M. R., M. D. Ohman, C. R. Benitez-Nelson, and M. R. Landry. 2013. Contributions of mesozooplankton to vertical carbon export in a coastal upwelling system. *Mar Ecol Prog Ser* **491**: 47-65.
- Sugie, K., H. Endo, K. Suzuki, J. Nishioka, H. Kiyosawa, and T. Yoshimura. 2013. Synergistic effects of pCO₂ and iron availability on nutrient consumption ratio of the Bering Sea phytoplankton community. *Biogeosciences* **10**: 6309-6321.
- Sunagawa, S., L. P. Coelho, S. Chaffron, J. R. Kultima, K. Labadie, G. Salazar, B. Djahanschiri, G. Zeller, D. R. Mende, A. Alberti, F. M. Cornejo-Castillo, P. I. Costea, C. Cruaud, F. Ovidio, S. Engelen, I. Ferrera, J. M. Gasol, L. Guidi, F. Hildebrand, F. Kokoszka, C. Lepoivre, G. Lima-Mendez, J. Poulain, B. T. Poulos, M. Royo-Lluch, H. Sarmiento, S. Vieira-Silva, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, C. Bowler, C. de Vargas, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, F. Not, H. Ogata, S. Pesant, S. Speich, L. Stemann, M. B. Sullivan, J. Weissenbach, P. Wincker, E. Karsenti, J. Raes, S. G. Acinas, and P. Bork. 2015. Structure and function of the global ocean microbiome. *Science* **348**: 1261359.

- Sunda, W. G., N. M. Price, and F. M. Morel. 2005. Trace metal ion buffers and their use in culture studies, p. 35-63. *Algal Culturing Techniques*.
- Sutak, R., H. Botebol, P.-L. Blaiseau, T. Léger, F.-Y. Bouget, J.-M. Camadro, and E. Lesuisse. 2012. A Comparative Study of Iron Uptake Mechanisms in Marine Microalgae: Iron Binding at the Cell Surface Is a Critical Step. *Plant Physiol* **160**: 2271-2284.
- Tang, D., and F. M. M. Morel. 2006. Distinguishing between cellular and Fe-oxide-associated trace elements in phytoplankton. *Mar Chem* **98**: 18-30.
- Tang, S., A. Lomsadze, and M. Borodovsky. 2015. Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res* **43**: e78.
- Till, C. P., J. R. Solomon, N. R. Cohen, R. H. Lampe, A. Marchetti, T. H. Coale, and K. W. Bruland. 2019. The iron limitation mosaic in the California Current System: Factors governing Fe availability in the shelf/near-shelf region. *Limnol Oceanogr* **64**: 109-123.
- Tortell, P. D., G. R. DiTullio, D. M. Sigman, and F. M. M. Morel. 2002. CO₂ effects on taxonomic composition and nutrient utilization in an Equatorial Pacific phytoplankton assemblage. *Mar Ecol Prog Ser* **236**: 37-43.
- Tortell, P. D., G. H. Rau, and F. M. M. Morel. 2000. Inorganic carbon acquisition in coastal Pacific phytoplankton communities. *Limnol Oceanogr* **45**: 1485-1500.
- Tortell, P. D., J. R. Reinfelder, and F. M. M. Morel. 1997. Active uptake of bicarbonate by diatoms. *Nature* **390**: 243-244.
- Trimborn, S., T. Brenneis, C. J. M. Hoppe, L. M. Laglera, L. Norman, J. Santos-Echeandía, C. Völkner, D. Wolf-Gladrow, and C. S. Hassler. 2017. Iron sources alter the response of Southern Ocean phytoplankton to ocean acidification. *Mar Ecol Prog Ser* **578**: 35-50.
- Trimborn, S., C. J. M. Hoppe, B. B. Taylor, A. Bracher, and C. Hassler. 2015. Physiological characteristics of open ocean and coastal phytoplankton communities of Western Antarctic Peninsula and Drake Passage waters. *Deep Sea Res Part I: Oceanogr Res Pap* **98**: 115-124.
- Tripathi, A. K., C. D. Roberts, and R. A. Eagle. 2009. Coupling of CO₂ and Ice Sheet Stability Over Major Climate Transitions of the Last 20 Million Years. *Science* **326**: 1394-1397.
- Turi, G., Z. Lachkar, and N. Gruber. 2014. Spatiotemporal variability and drivers of pCO₂ and air-sea CO₂ fluxes in the California Current System: an eddy-resolving modeling study. *Biogeosciences* **11**: 671-690.
- Turnšek, J., J. K. Brunson, M. d. P. M. Viedma, T. J. Deerinck, A. Horák, M. Oborník, V. A. Bielinski, and A. E. Allen. 2021. Proximity proteomics in a marine diatom reveals a putative cell surface-to-chloroplast iron trafficking pathway. *eLife* **10**: e52770.
- Uppstrom, L. 1974. The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. *Deep Sea Res.* **21**: 161-162.

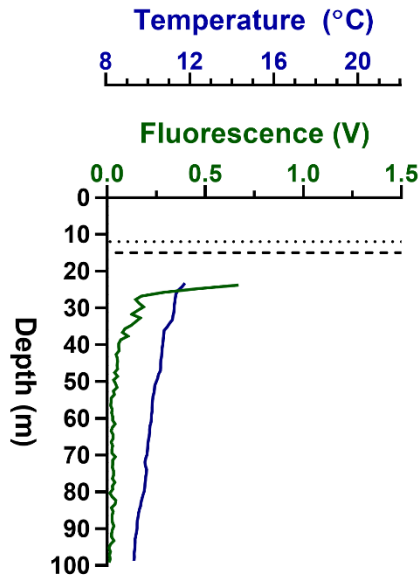
- Urzica, E. I., D. Casero, H. Yamasaki, S. I. Hsieh, L. N. Adler, S. J. Karpowicz, C. E. Blaby-Haas, S. G. Clarke, J. A. Loo, M. Pellegrini, and S. S. Merchant. 2012. Systems and Trans-System Level Analysis Identifies Conserved Iron Deficiency Responses in the Plant Lineage. *Plant Cell* **24**: 3921-3948.
- Valenzuela, J. J., A. López García de Lomana, A. Lee, E. V. Armbrust, M. V. Orellana, and N. S. Baliga. 2018. Ocean acidification conditions increase resilience of marine diatoms. *Nat Commun* **9**: 2328.
- Vargas, C. A., N. A. Lagos, M. A. Lardies, C. Duarte, P. H. Manríquez, V. M. Aguilera, B. Broitman, S. Widdicombe, and S. Dupont. 2017. Species-specific responses to ocean acidification should account for local adaptation and adaptive plasticity. *Nat Ecol Evol* **1**: 0084.
- Wang, X. J., M. Behrenfeld, R. Le Borgne, R. Murtugudde, and E. Boss. 2009. Regulation of phytoplankton carbon to chlorophyll ratio by light, nutrients and temperature in the Equatorial Pacific Ocean: a basin-scale model. *Biogeosciences* **6**: 391-404.
- Weiss, S., Z. Z. Xu, S. Peddada, A. Amir, K. Bittinger, A. Gonzalez, C. Lozupone, J. R. Zaneveld, Y. Vázquez-Baeza, A. Birmingham, E. R. Hyde, and R. Knight. 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5**: 27.
- Weisser, H., and J. S. Choudhary. 2017. Targeted feature detection for data-dependent shotgun proteomics. *J Proteome Res* **16**: 2964-2974.
- Worden, A. Z., J. Janouskovec, D. McRose, A. Engman, R. M. Welsh, S. Malfatti, S. G. Tringe, and P. J. Keeling. 2012. Global distribution of a wild alga revealed by targeted metagenomics. *Curr Biol* **22**: R675-R677.
- Yoshimura, T., K. Sugie, H. Endo, K. Suzuki, J. Nishioka, and T. Ono. 2014. Organic matter production response to CO₂ increase in open subarctic plankton communities: Comparison of six microcosm experiments under iron-limited and -enriched bloom conditions. *Deep Sea Res Part I: Oceanogr Res Pap* **94**: 1-14.
- Yoshimura, T., K. Suzuki, H. Kiyosawa, T. Ono, H. Hattori, K. Kuma, and J. Nishioka. 2013. Impacts of elevated CO₂ on particulate and dissolved organic matter production: microcosm experiments using iron-deficient plankton communities in open subarctic waters. *J Oceanogr* **69**: 601-618.
- Zhu, K., M. J. Hopwood, J. E. Groenenberg, A. Engel, E. P. Achterberg, and M. Gledhill. 2021. Influence of pH and Dissolved Organic Matter on Iron Speciation and Apparent Iron Solubility in the Peruvian Shelf and Slope Region. *Environ Sci & Technol* **55**: 9372-9383.

CHAPTER 2 SUPPLEMENTAL INFORMATION

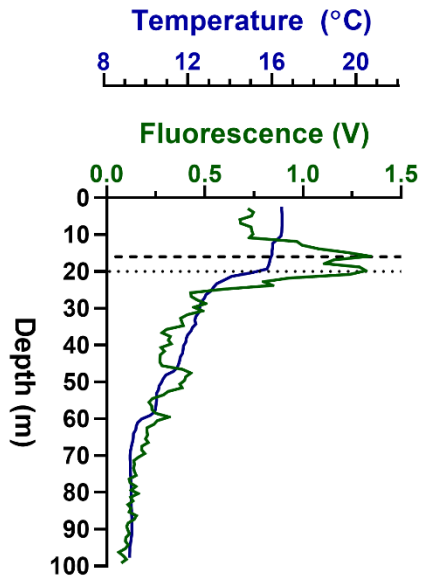
E1



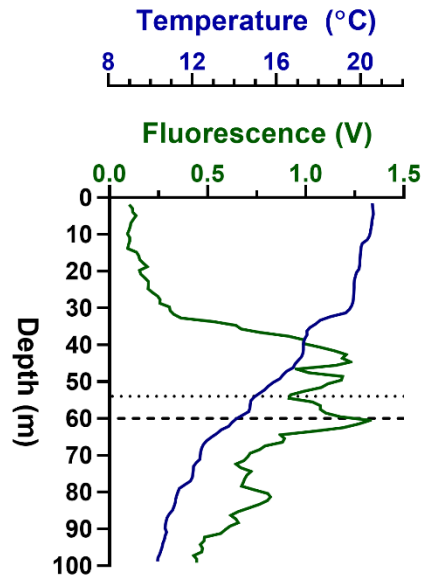
E2



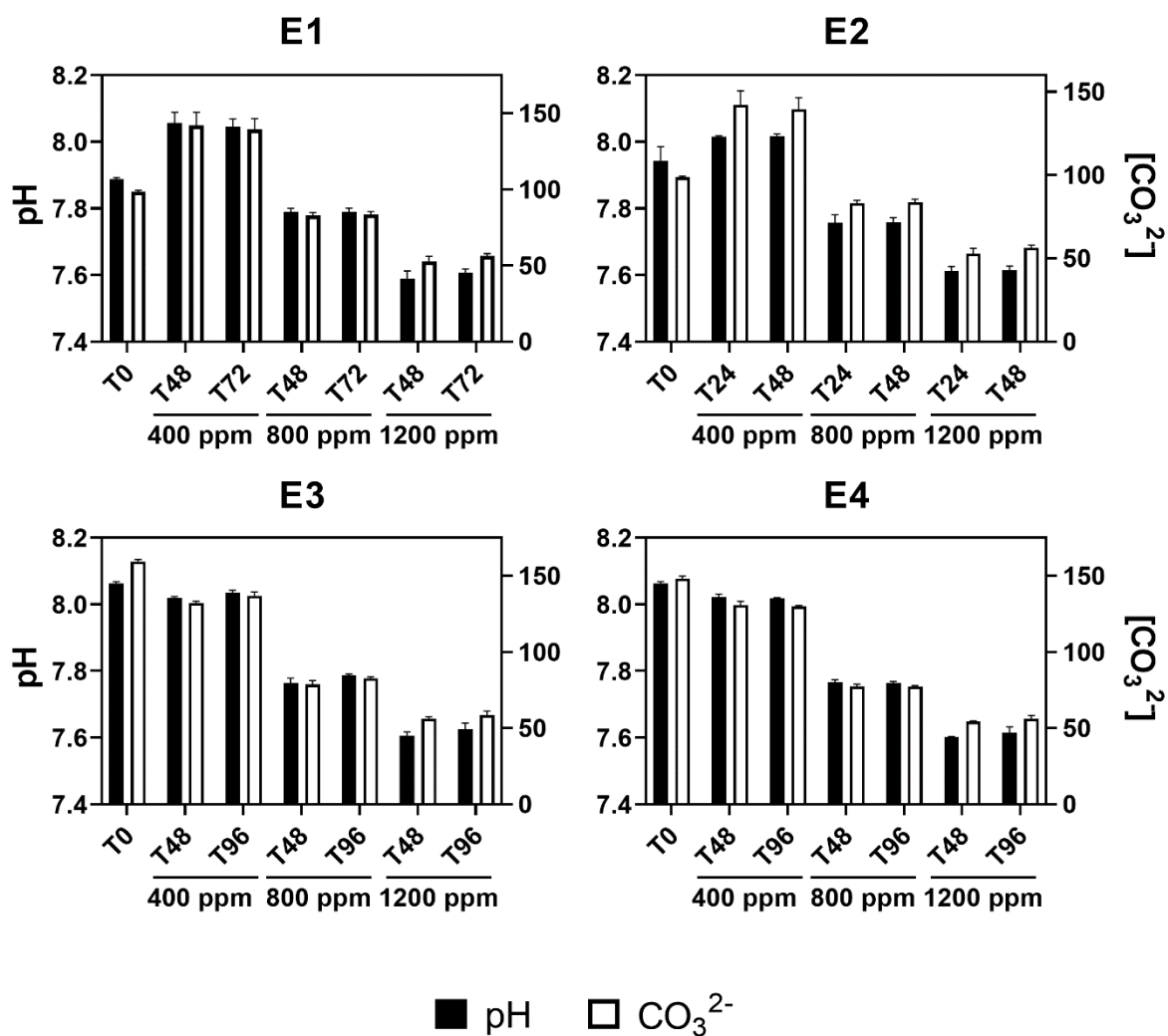
E3



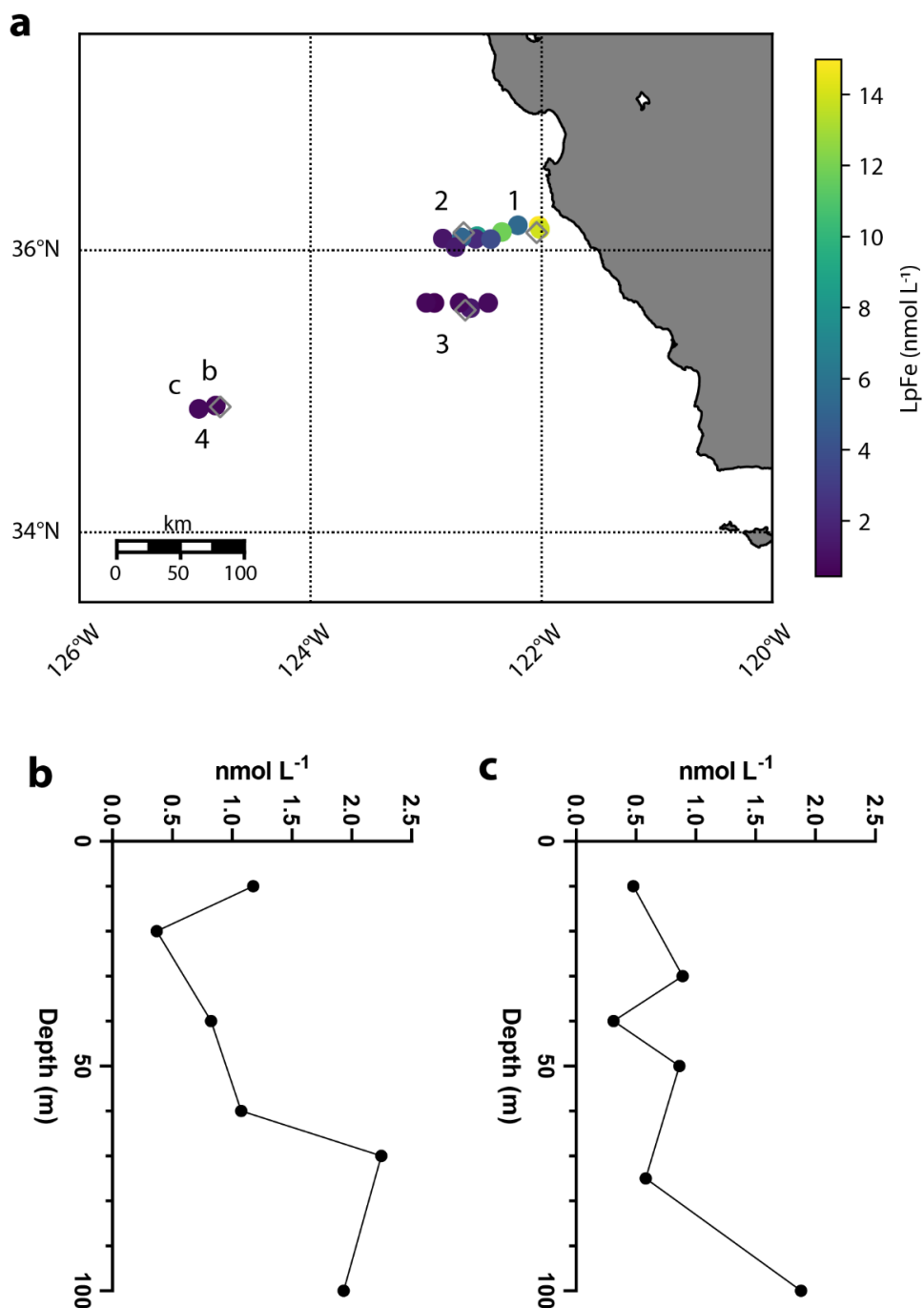
E4



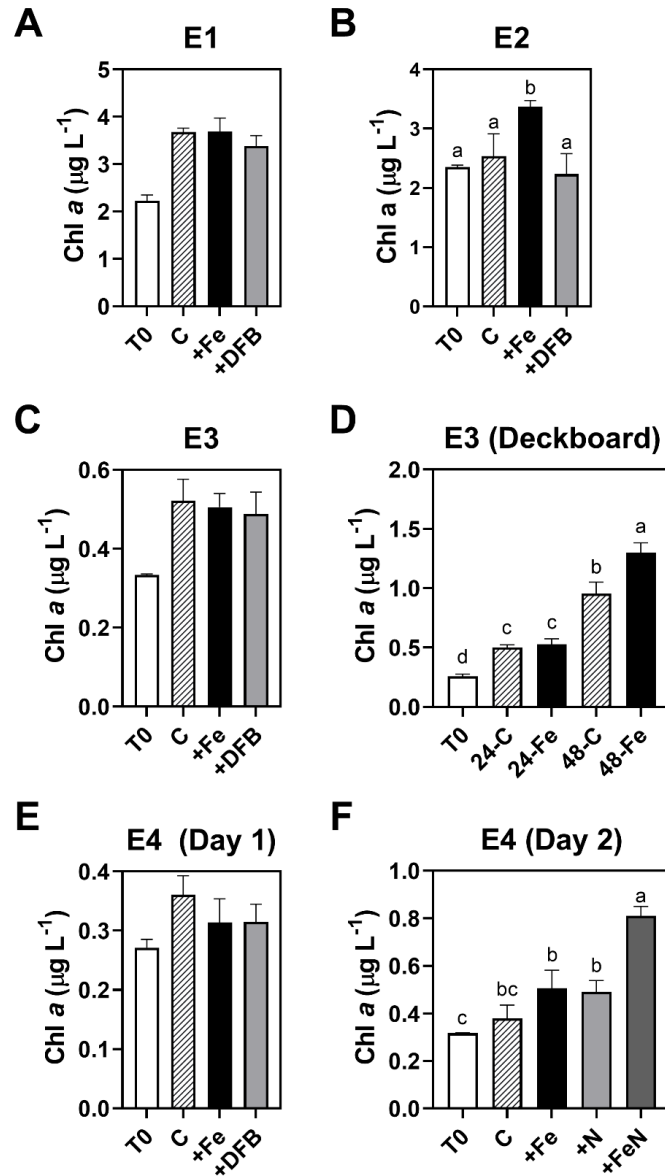
Supplemental Figure 2.1: Temperature (°C, blue) and chlorophyll fluorescence voltages (V, green) from the CTD downcast prior to the start of the incubations. The horizontal dashed line indicates the depth of seawater collection for the ocean acidification experiments. The dotted line indicates the depth of collection for the *in situ* incubations. Data shallower than 24 m are not available for the cast for Experiment 2.



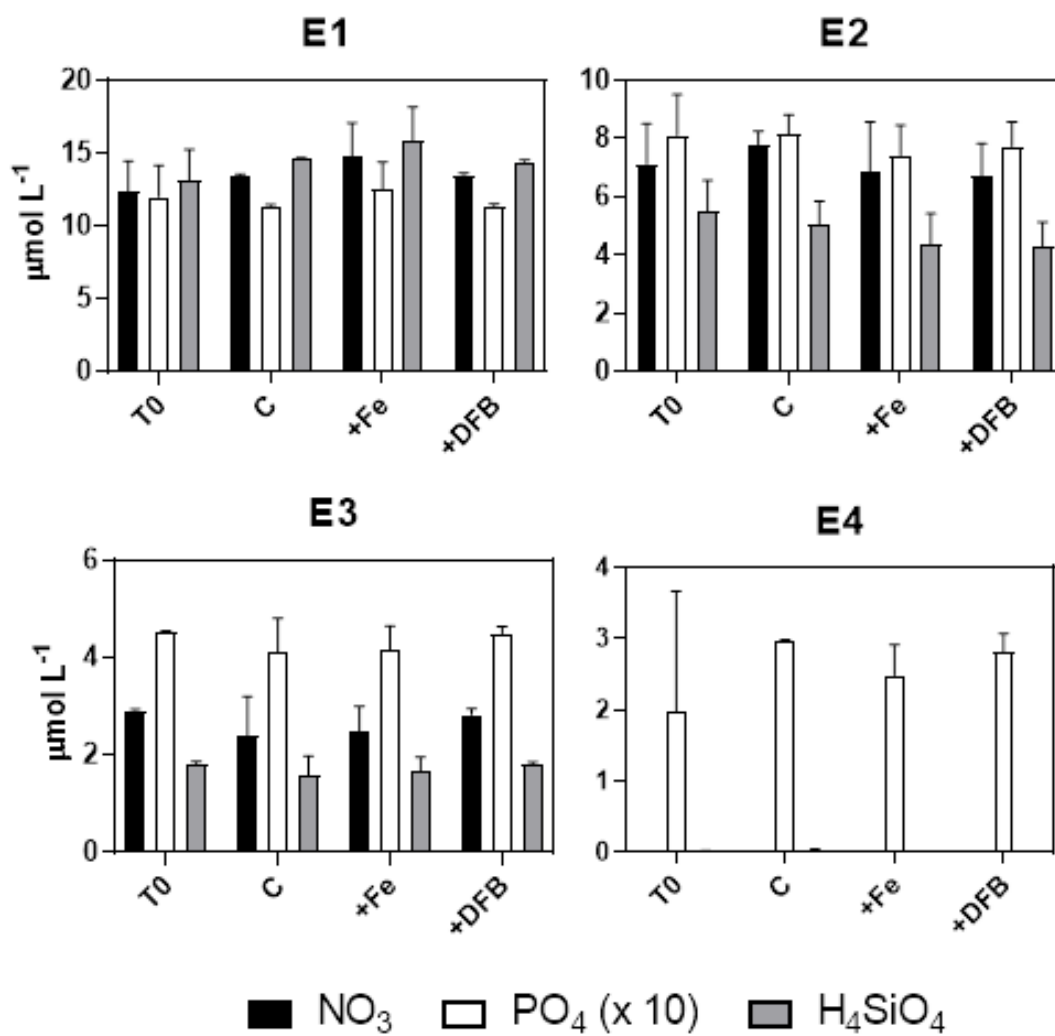
Supplemental Figure 2.2: pH (total scale, left y-axis) and carbonate ion concentrations [CO₃²⁻] (μmol kg⁻¹, right y-axis) computed from dissolved inorganic carbon (DIC) and total alkalinity (TA) for each experiment (E). Timepoints (T) and CO₂ treatments (μatm) are denoted on the x-axis. Measured DIC and TA as well as calculated pH, pCO₂, and [CO₃²⁻] are shown in Supplemental Table 2.3.



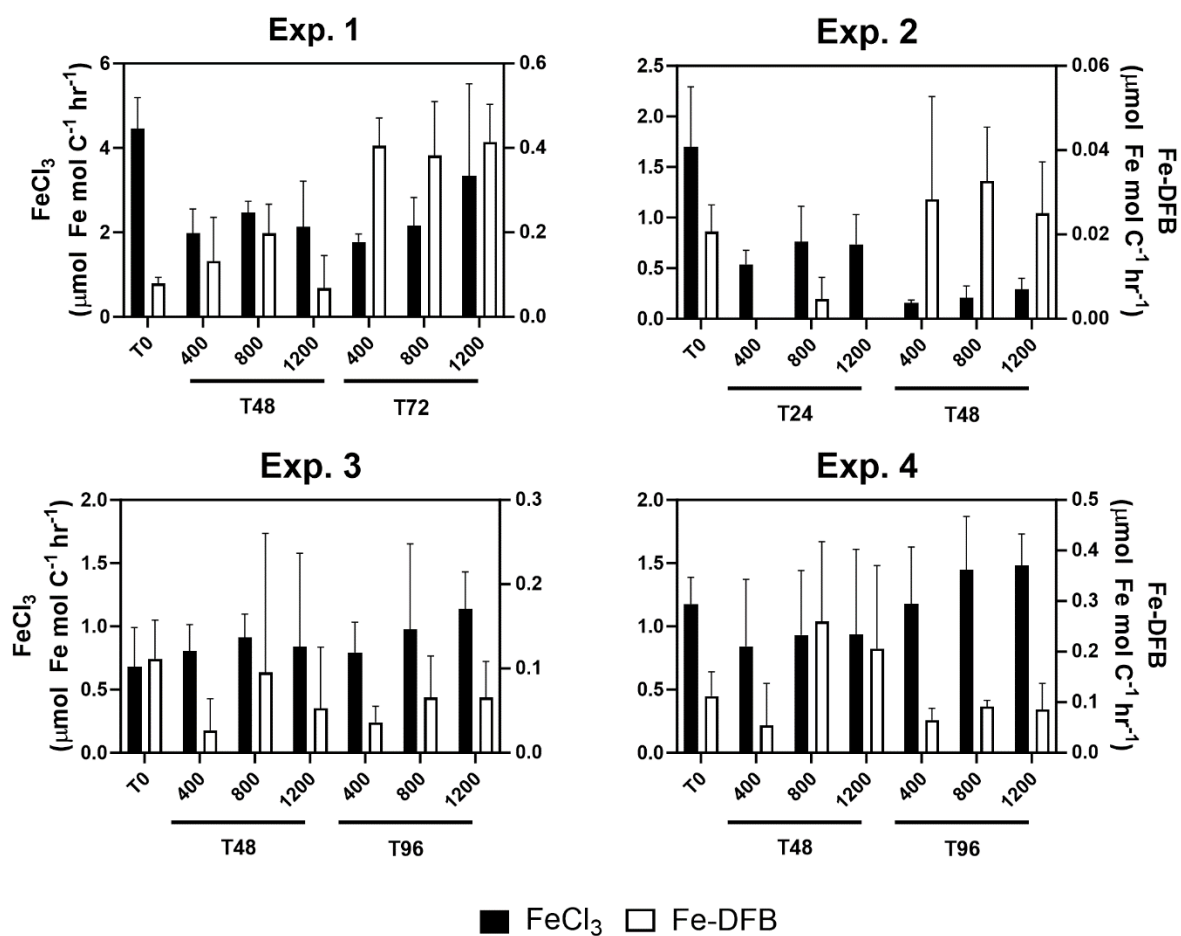
Supplemental Figure 2.3: Labile particulate iron concentrations (LpFe, nmol L⁻¹) calculated from total dissolvable iron minus dissolved iron concentrations. (a) Average concentrations from 10 to 25 m at locations near where the ocean acidification experiments were conducted (n = 1-2). Locations where seawater was collected for each experiment are numbered and shown as grey diamonds. (b, c) Depth profiles corresponding to the sites labelled b and c in panel a near the initial seawater collection for Experiment 4 at 60 m.



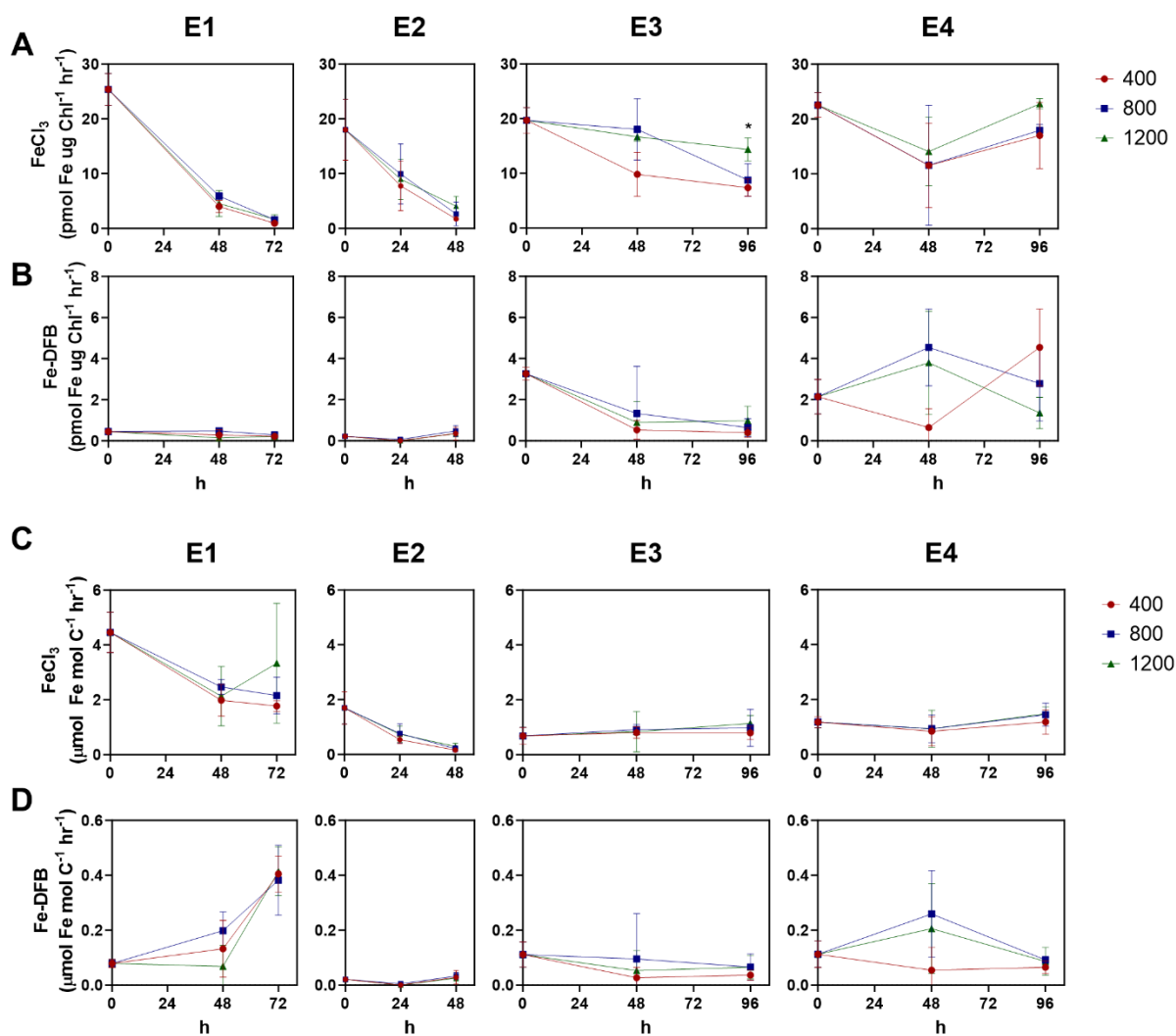
Supplemental Figure 2.4: Chlorophyll a concentrations ($\mu\text{g L}^{-1}$) from the parallel incubations. (A, B, C, E) In situ experiments with control, iron addition (+Fe), or DFB addition (+DFB) treatments after 24 hours. (D) Deckboard experiment during Experiment 3 with control and iron addition treatments sampled after 24 and 48 hours. (F) In situ incubation initiated 24 hours after Experiment 4 with control, iron addition, nitrate addition, and iron plus nitrate addition treatments. Bottles were incubated for 48 hours. Y-axis scales are different for each experiment. Letters denote significant differences between means where observed ($P < 0.05$, ANOVA with Tukey Test). Error bars represent the standard deviation of the mean ($n = 3$).



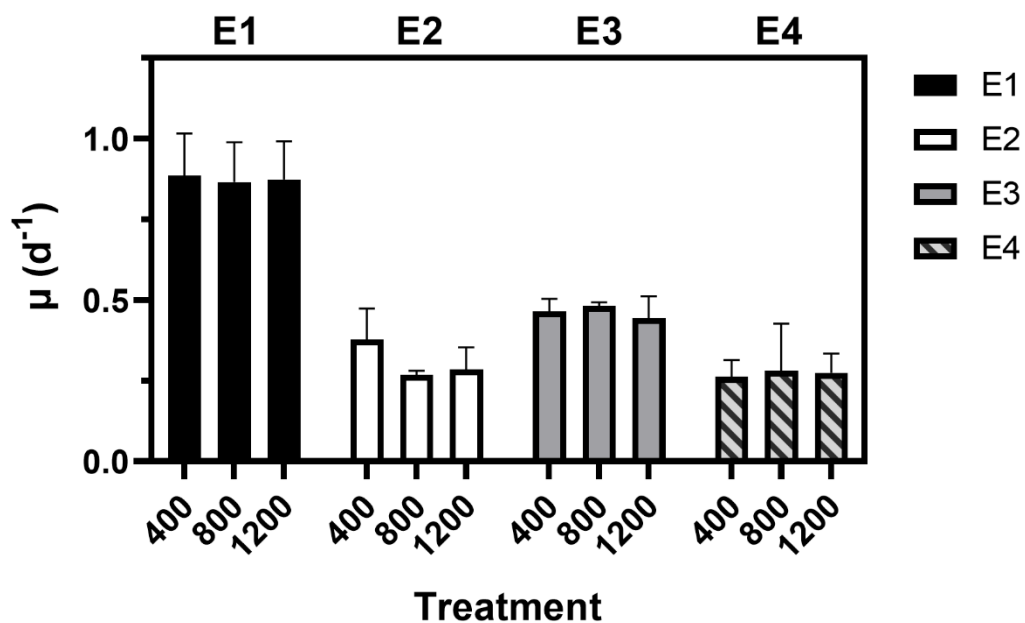
Supplemental Figure 2.5: Macronutrient concentrations ($\mu\text{mol L}^{-1}$) from the parallel *in situ* incubations at the initial time point (T0) and after 24 hours of incubation with control (C), iron addition (+Fe), or DFB addition (+DFB) treatments. Phosphate (PO_4) concentrations are one-tenth the plotted values. Y-axis scales are different for each experiment. Error bars represent the standard deviation of the mean ($n = 3$).



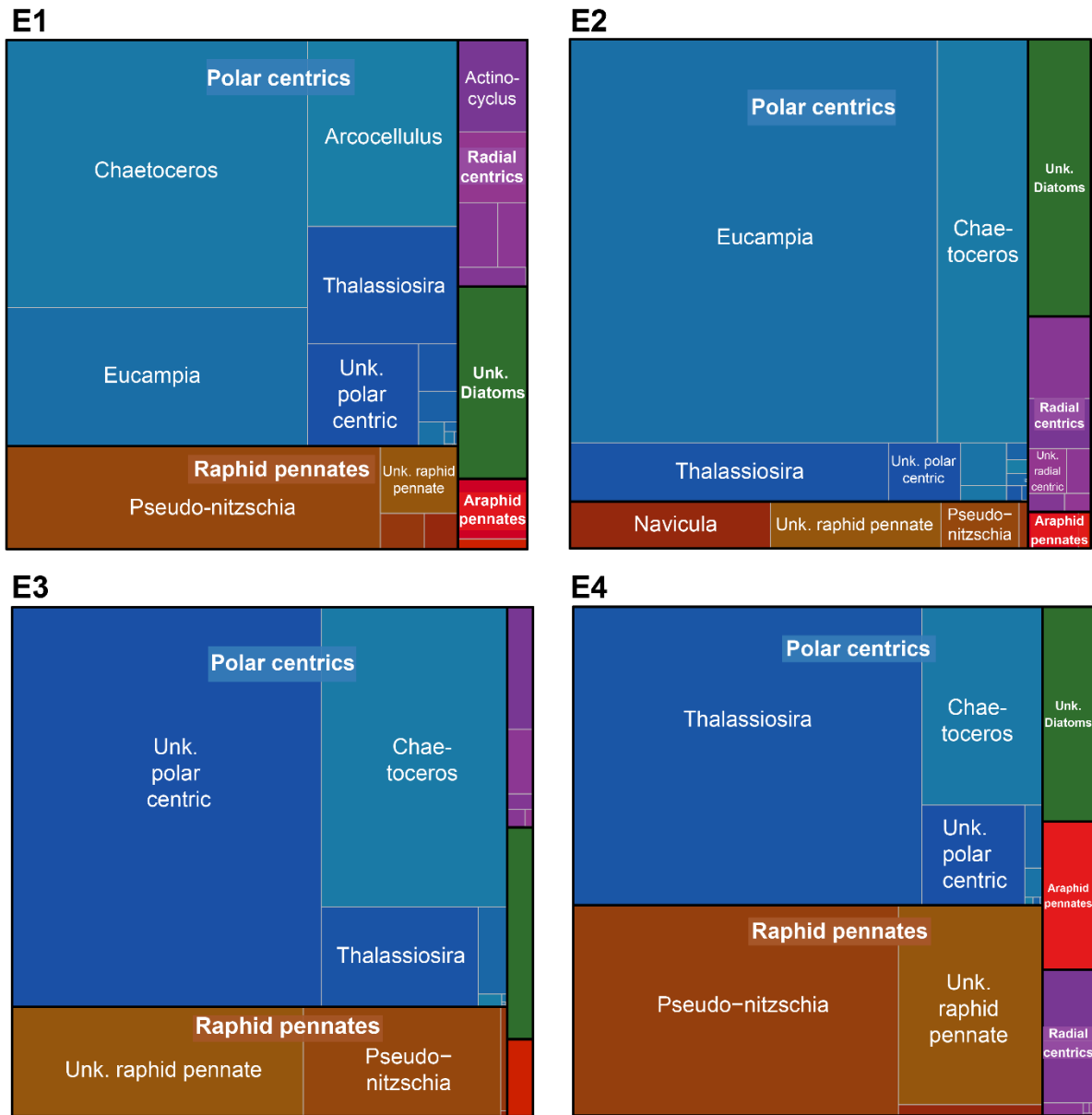
Supplemental Figure 2.6: Short-term iron-59 uptake rates normalized to particulate organic carbon ($\mu\text{mol Fe mol C}^{-1} \text{ hr}^{-1}$) for inorganic iron (FeCl_3 , black, left y-axis) and organically-complexed iron as ferrioxamine B (FeDFB, white, right y-axis). Error bars represent the standard deviation of the mean ($n = 3$ except the Experiment 4 800 ppm samples where $n = 2$).



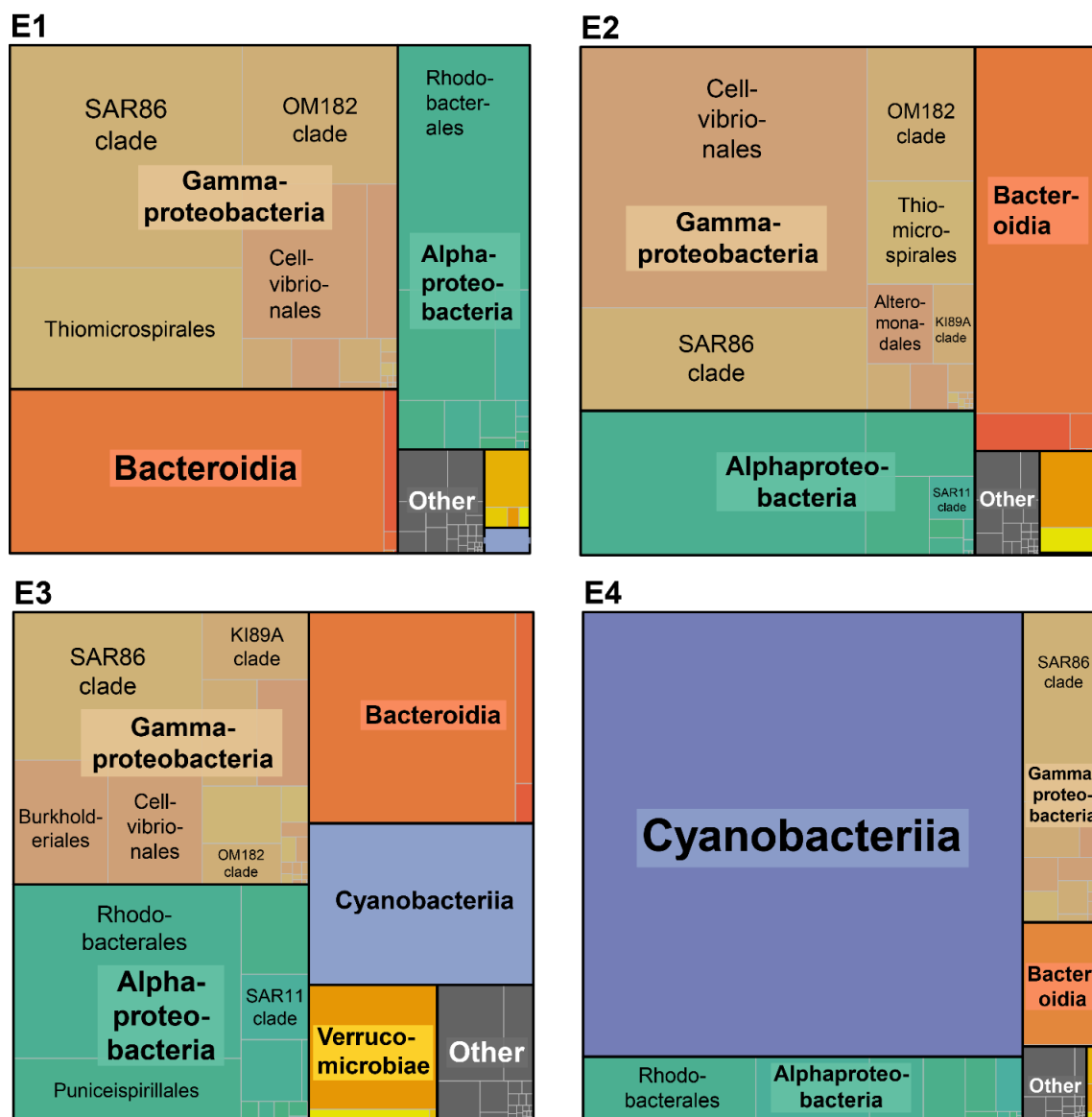
Supplemental Figure 2.7: Short-term iron-59 uptake rates over time. (A, B) normalized to chlorophyll *a* concentrations (pmol Fe μg Chl⁻¹ hr⁻¹). (C, D) normalized to particulate organic carbon (μmol Fe mol C⁻¹ hr⁻¹). (A, C) inorganic iron (FeCl₃) uptakes. (B, D) organically-complexed iron as ferrioxamine B (FeDFB) uptake rates. Error bars represent the standard deviation of the mean ($n = 3$ except the Experiment 4 800 ppm samples where $n = 2$).



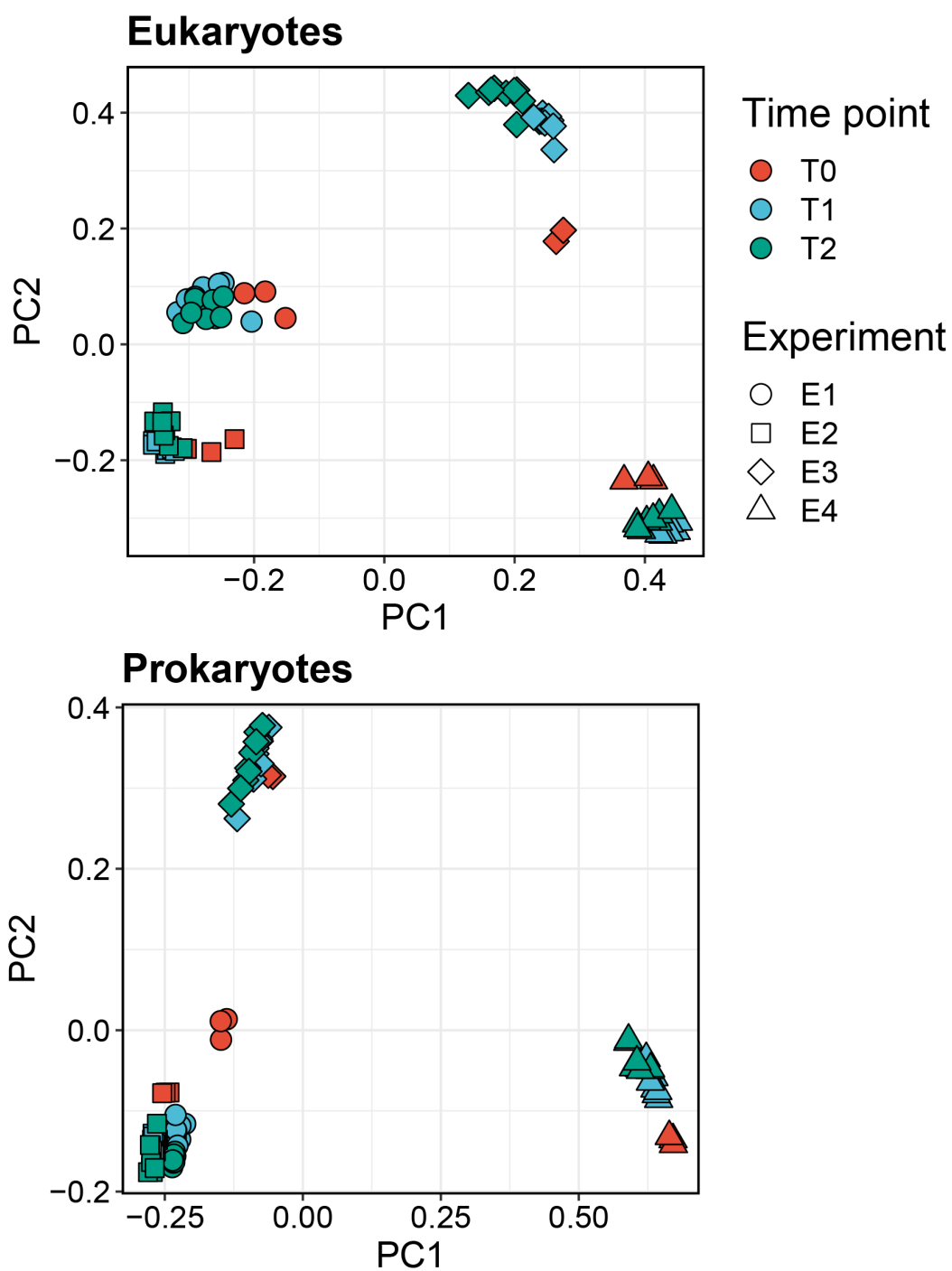
Supplemental Figure 2.8: Community growth rates (μ , d^{-1}) for each treatment within each experiment. Growth rates were not significantly different within each experiment ($P > 0.05$).



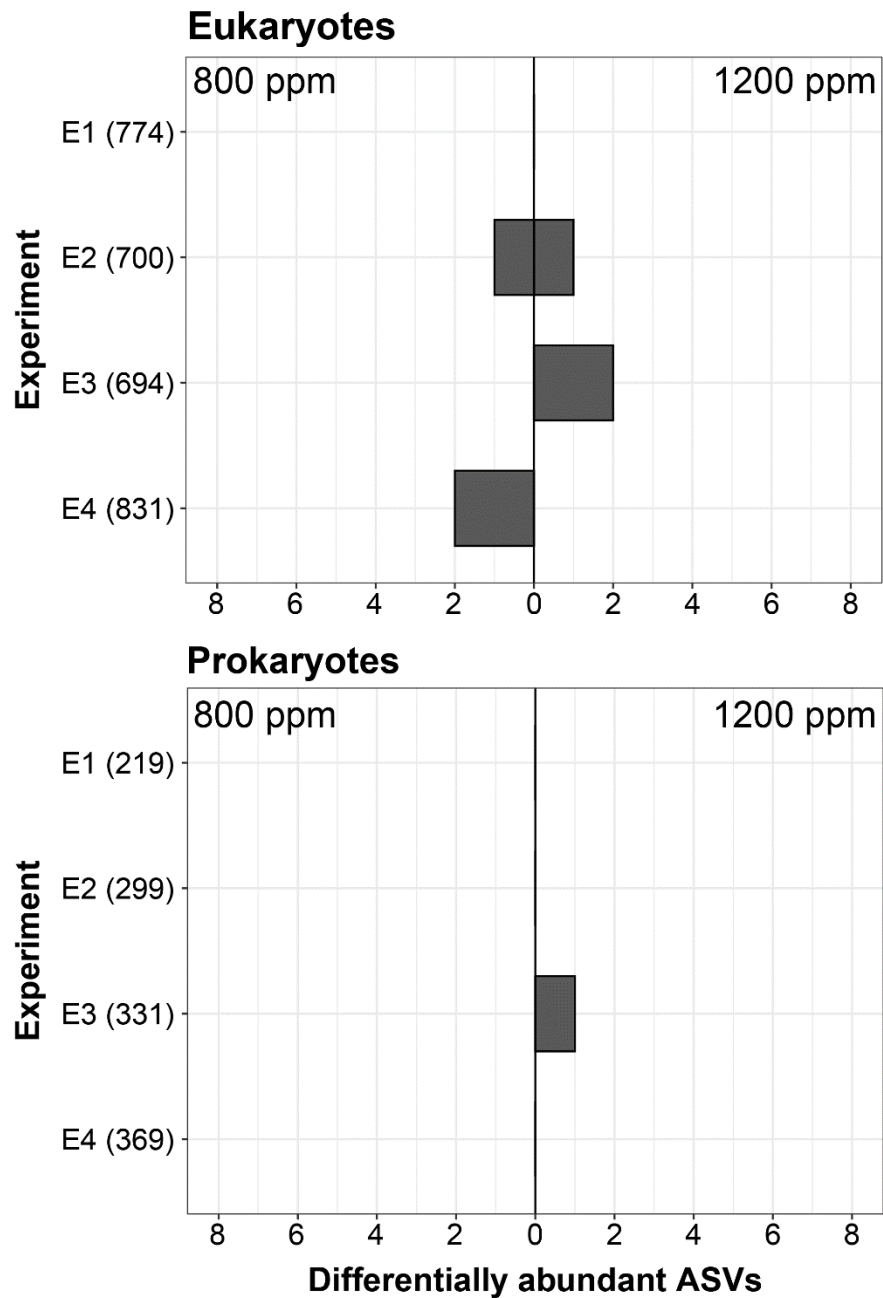
Supplemental Figure 2.9: The relative abundance of diatom orders and genera within diatom annotated 18S rRNA amplicons at each experiment (E1-4). Orders share colours as follows: polar centric (blue), radial centric (purple), raphid pennate (brown), araphid pennate (red), and unknown (green).



Supplemental Figure 2.10: The relative abundance of bacterial classes and orders from 16S rRNA amplicons at each experiment (E1-4).

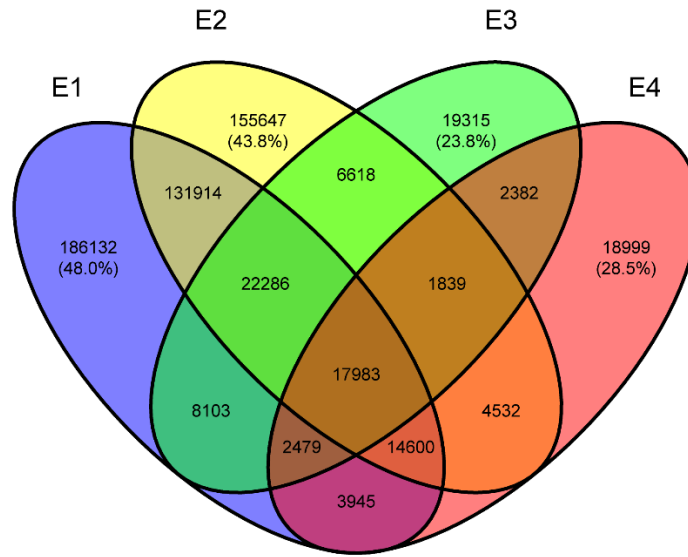


Supplemental Figure 2.11: PCoA plots of Bray-Curtis dissimilarity for the eukaryotic community (18S rRNA, top) and prokaryotic community (16S rRNA, bottom). Experiments and time points are denoted by colour and shape respectively (see legend).

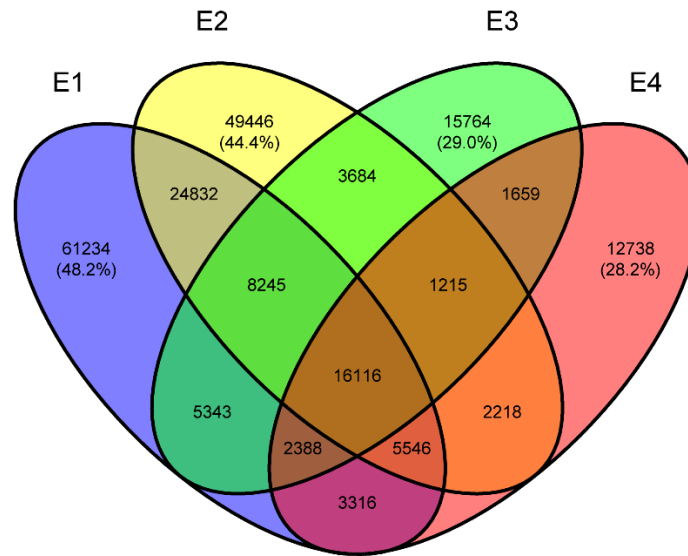


Supplemental Figure 2.12: The number of significantly differential abundant amplicon sequence variants (ASVs) at each experiment determined by DESeq2 between the 800 μ atm (right) and 400 μ atm (left) treatments at the final time point for each experiment ($P < 0.05$). Eukaryotes are shown on top and prokaryotes are shown on the bottom plot. The total number of ASVs for each experiment is shown in parentheses next to the experiment number of the y-axis.

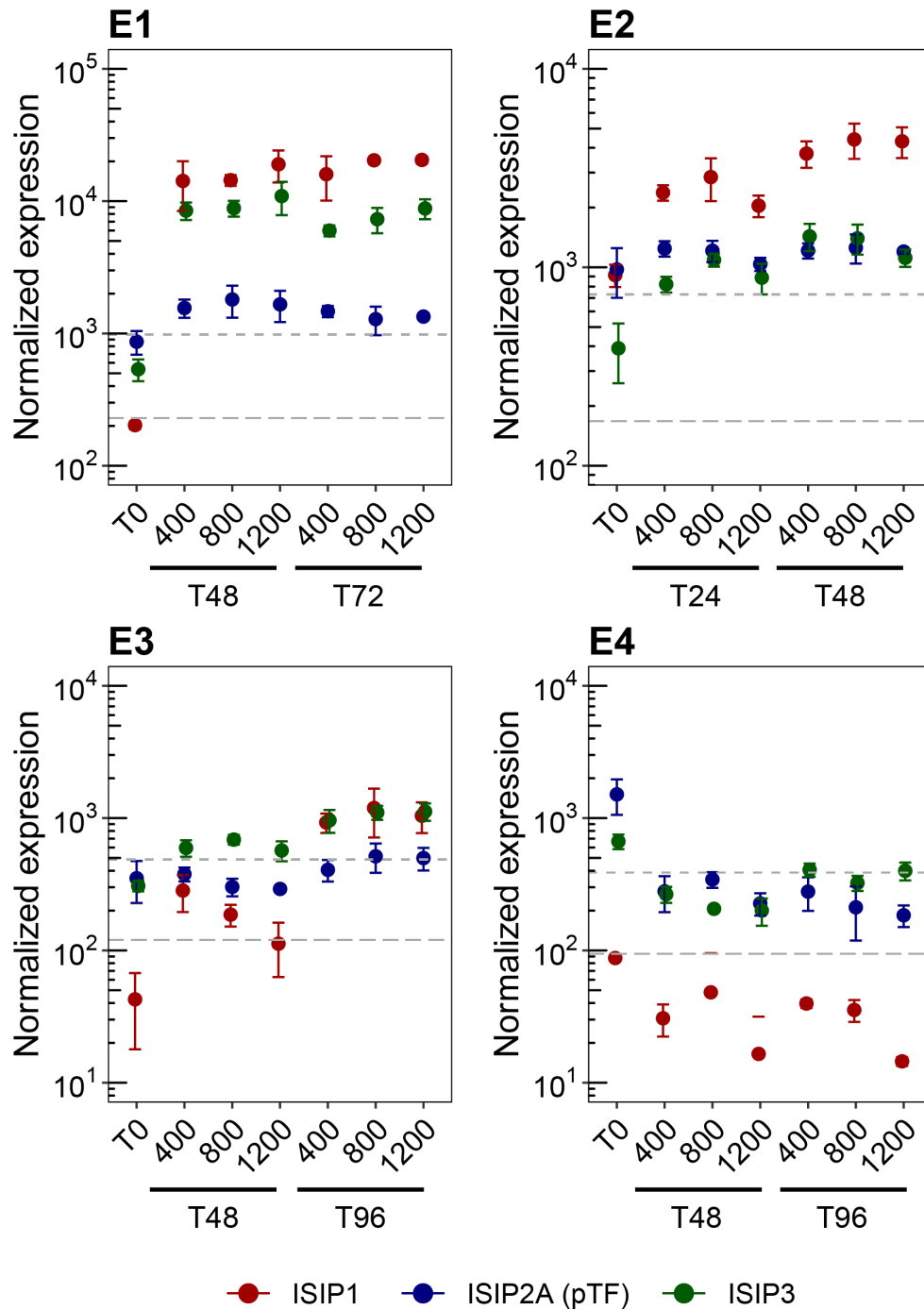
A. Centric Diatoms



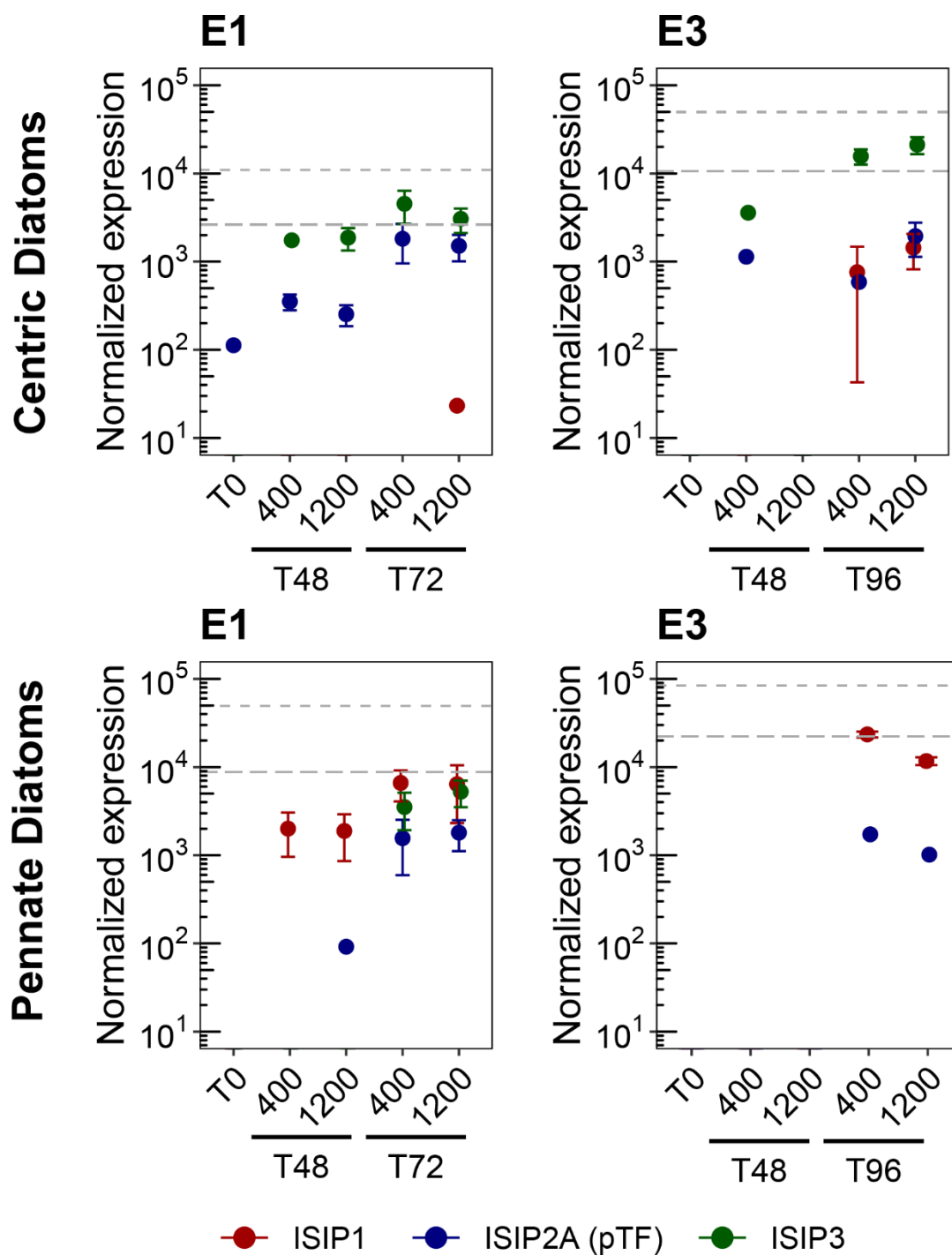
B. Pennate Diatoms



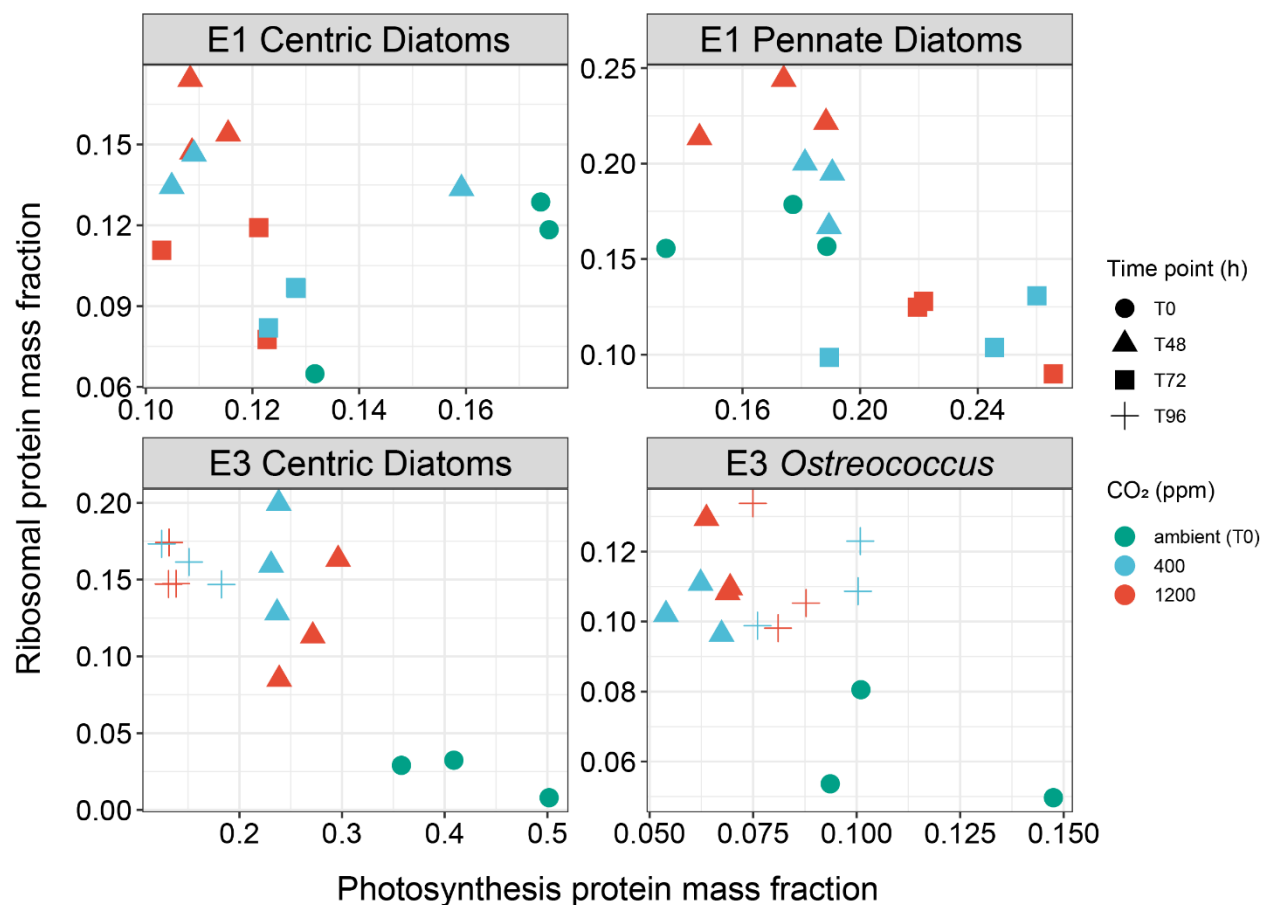
Supplemental Figure 2.13: Venn diagrams of detected orthologous genes among the communities at the different experiments (E1-E4) for (A) centric diatoms and (B) pennate diatoms. The percentage of orthologous genes unique to each experiment is shown.



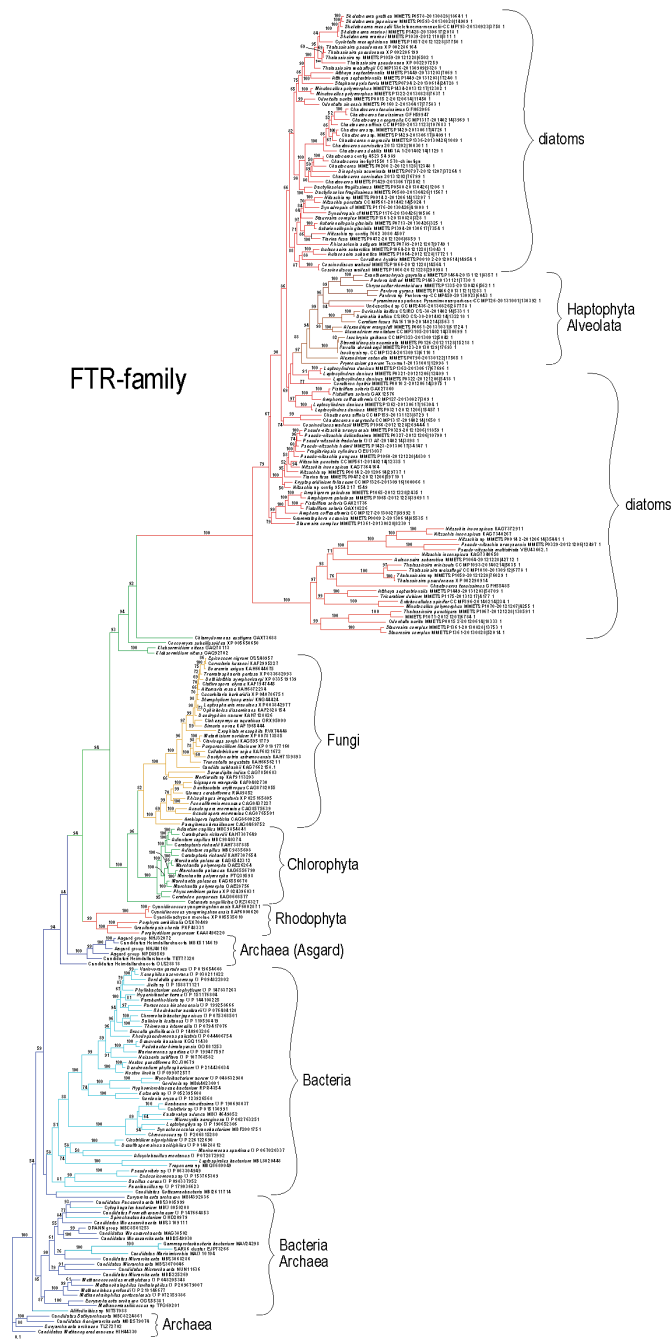
Supplemental Figure 2.14: Normalized transcript expression of iron starvation-induced proteins (ISIPs) in pennate diatoms (log₁₀ scale): ISIP1 (red), ISIP2A (phytoferritin or pTF, blue), ISIP3 (green). The horizontal dashed lines indicate the averaged 95th and 99th percentile for transcript abundance in each experiment. Error bars represent the standard deviation of the mean (n = 3 except the Experiment 3 T0 and Experiment 4 800 ppm samples where n = 2).



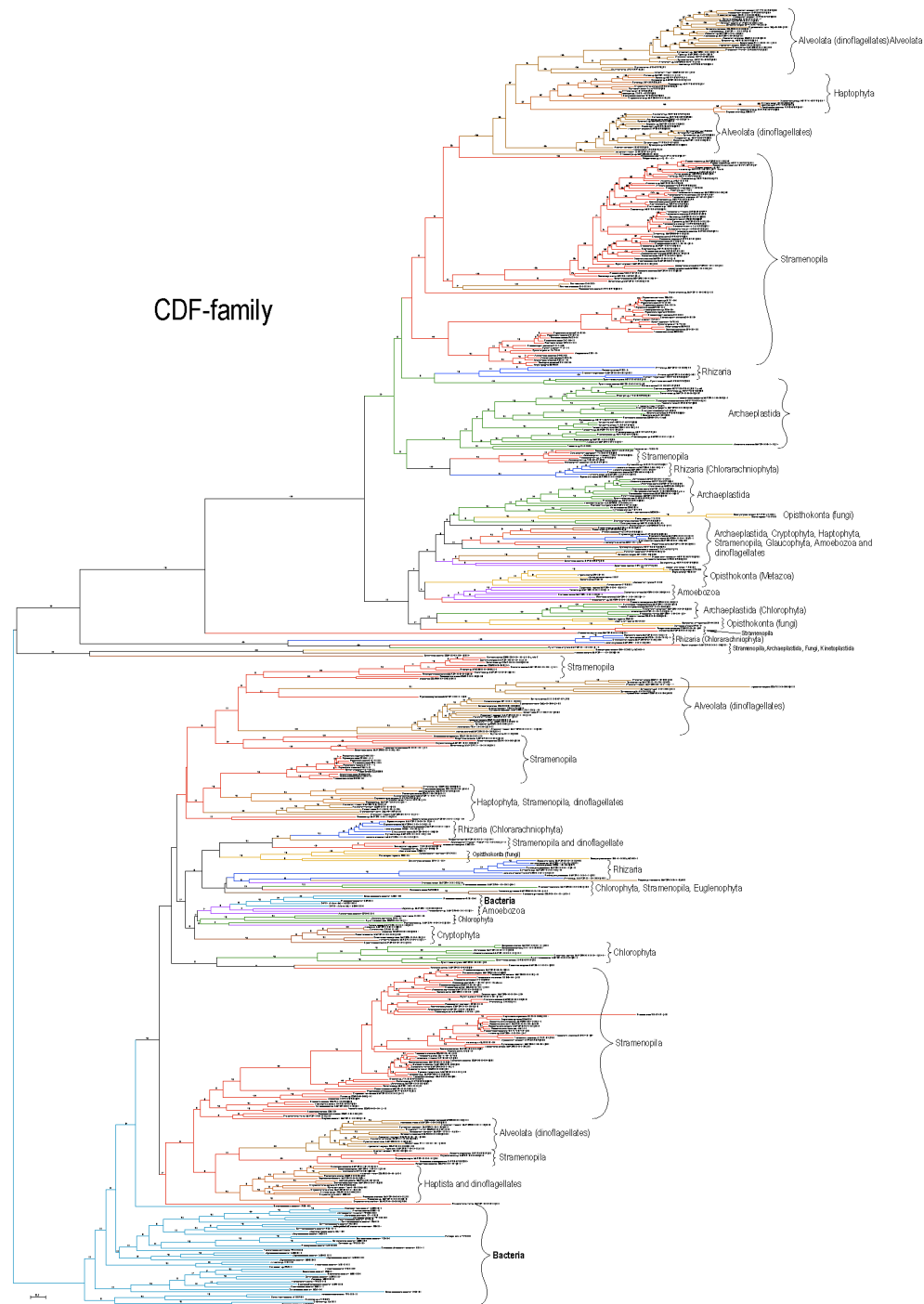
Supplemental Figure 2.15: Normalized (total sum scaled) protein abundances of iron starvation-induced proteins (ISIPs) in centric diatoms (top row) and pennate diatoms (bottom row). The horizontal dashed lines indicate the averaged 95th and 99th percentile for protein abundances in each experiment. Error bars represent the standard deviation of the mean (n = 3 except the Experiment 3 T0 samples where n = 2).



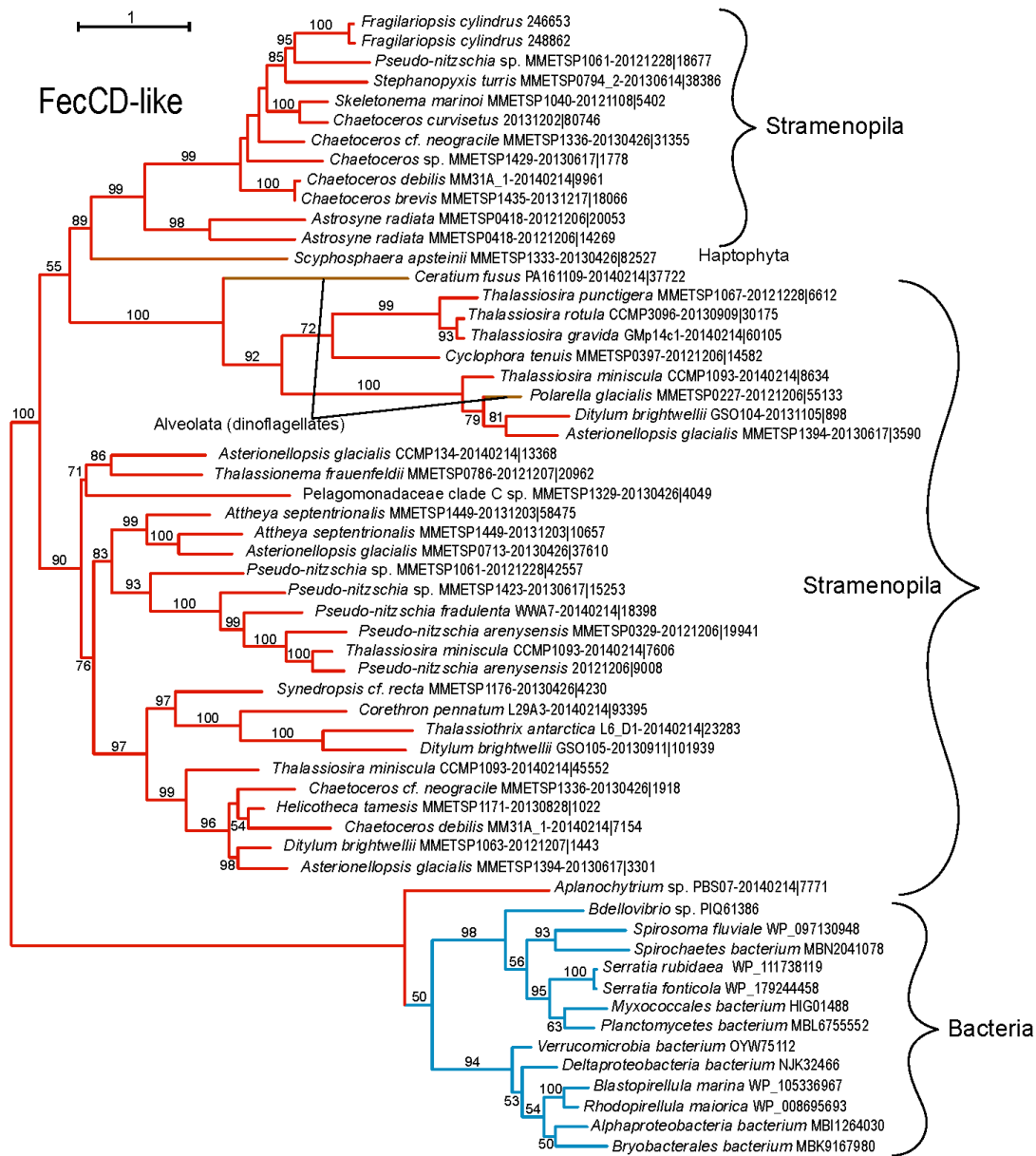
Supplemental Figure 2.16: Photosynthetic and ribosomal protein mass fractions for each Experiment (E), and taxonomic group with at least 50 unique peptides for each functional category. Time points (h) are denoted by shape and CO₂ treatments are denoted by color as shown in the legend. Note that the scale of each axis is different.



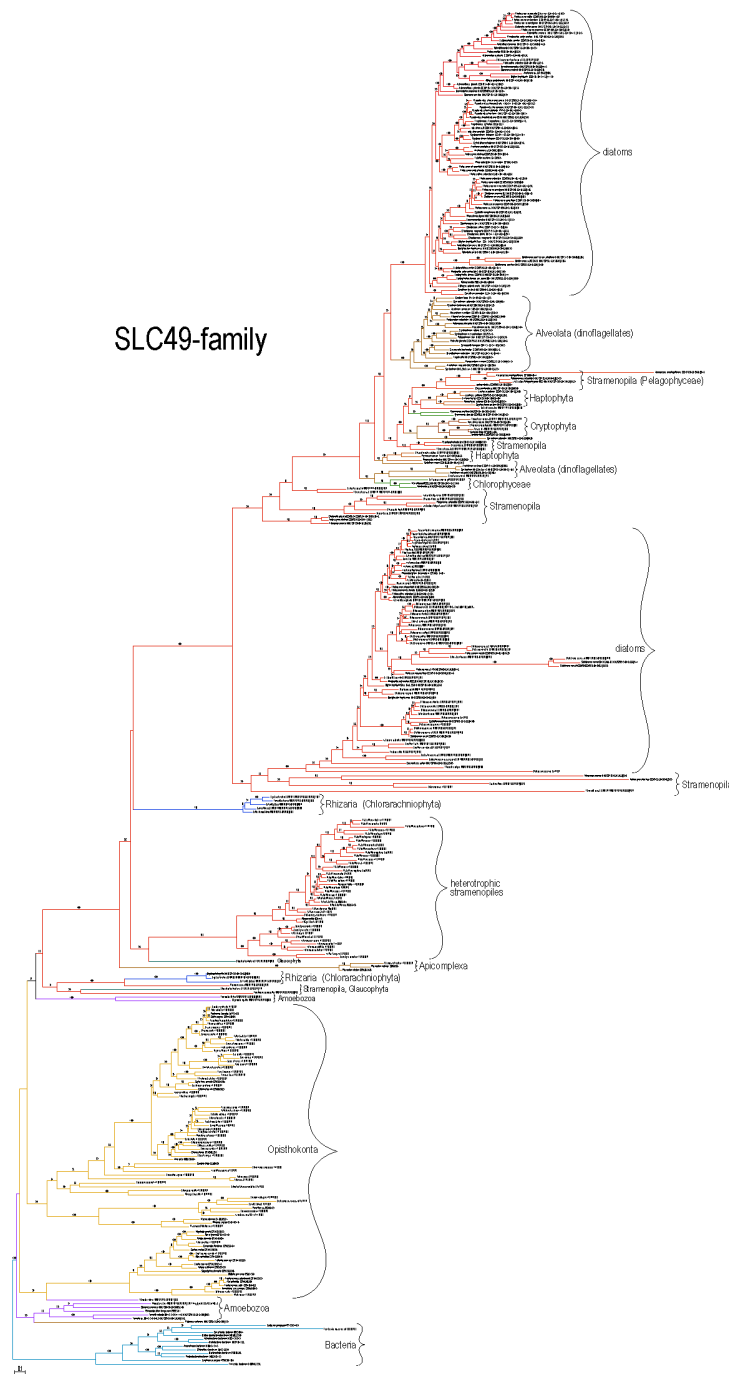
Supplemental Figure 2.17: Maximum likelihood phylogenetic tree as inferred from iron permease (FTR1) amino acid sequences (255 taxa/241 positions) under the LG+C40+F+R8 model as implemented in IQTree 2. The scale bar represents 0.1 amino acid substitutions per position. Branch colors and labels correspond to different taxonomic groups. Numbers above branches indicate ultra-fast Maximum-likelihood bootstrap support. A PDF version of this tree is available in this manuscript's associated figshare repository: doi.org/10.6084/m9.figshare.22589218



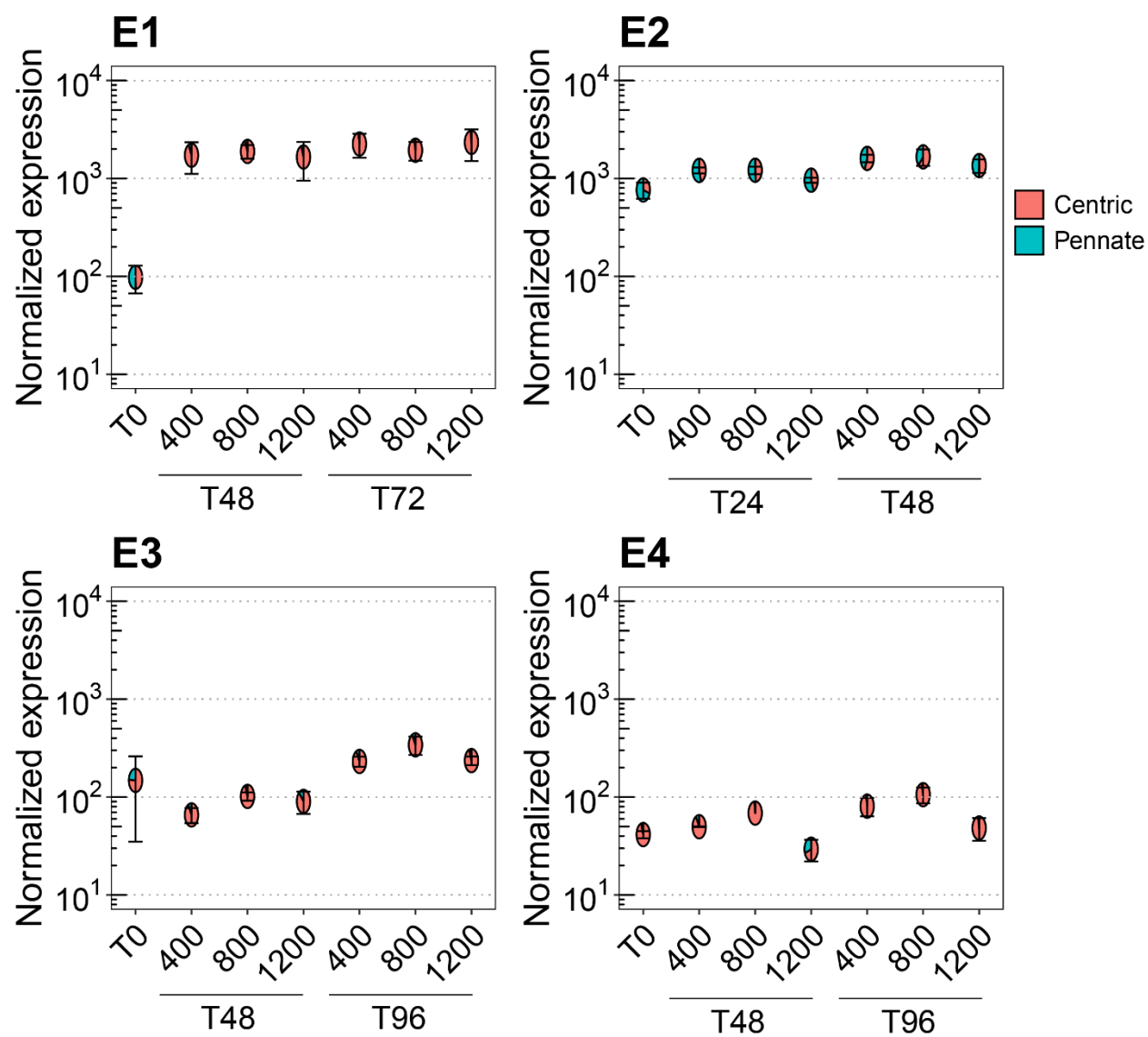
Supplemental Figure 2.18: Maximum likelihood phylogenetic tree as inferred from CDF-family amino acid sequences (542 taxa/309 positions) under the LG+F+R10 model as implemented in IQTree 2. The scale bar represents 0.1 amino acid substitutions per position. Branch colors and labels correspond to different taxonomic groups. Numbers above branches indicate Maximum likelihood ultra-fast bootstrap support. A PDF version of this tree is available in this manuscript's associated figshare repository: doi.org/10.6084/m9.figshare.22589218



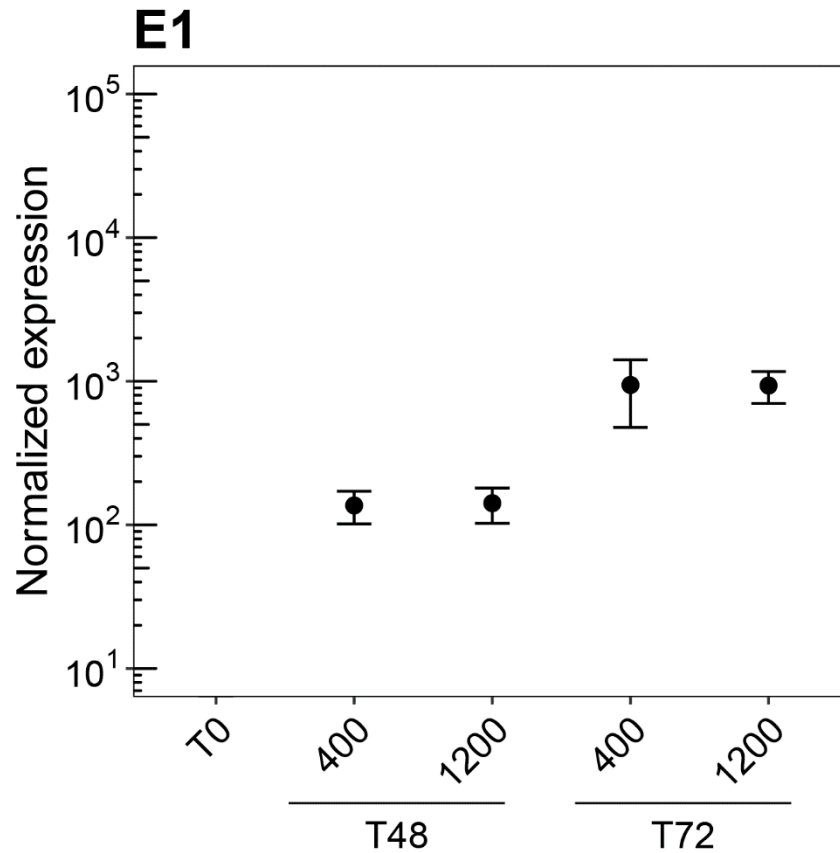
Supplemental Figure 2.19: Maximum likelihood phylogenetic tree as inferred from FecCD-like sequences (58 taxa/250 positions) under the LG+C20+F+R5 model as implemented in IQTree 2. The scale bar represents 0.1 amino acid substitutions per position. Branch colours and labels correspond to different taxonomic groups. Numbers above branches indicate Maximum likelihood ultra-fast bootstrap support. A PDF version of this tree is available in this manuscript's associated figshare repository: doi.org/10.6084/m9.figshare.22589218



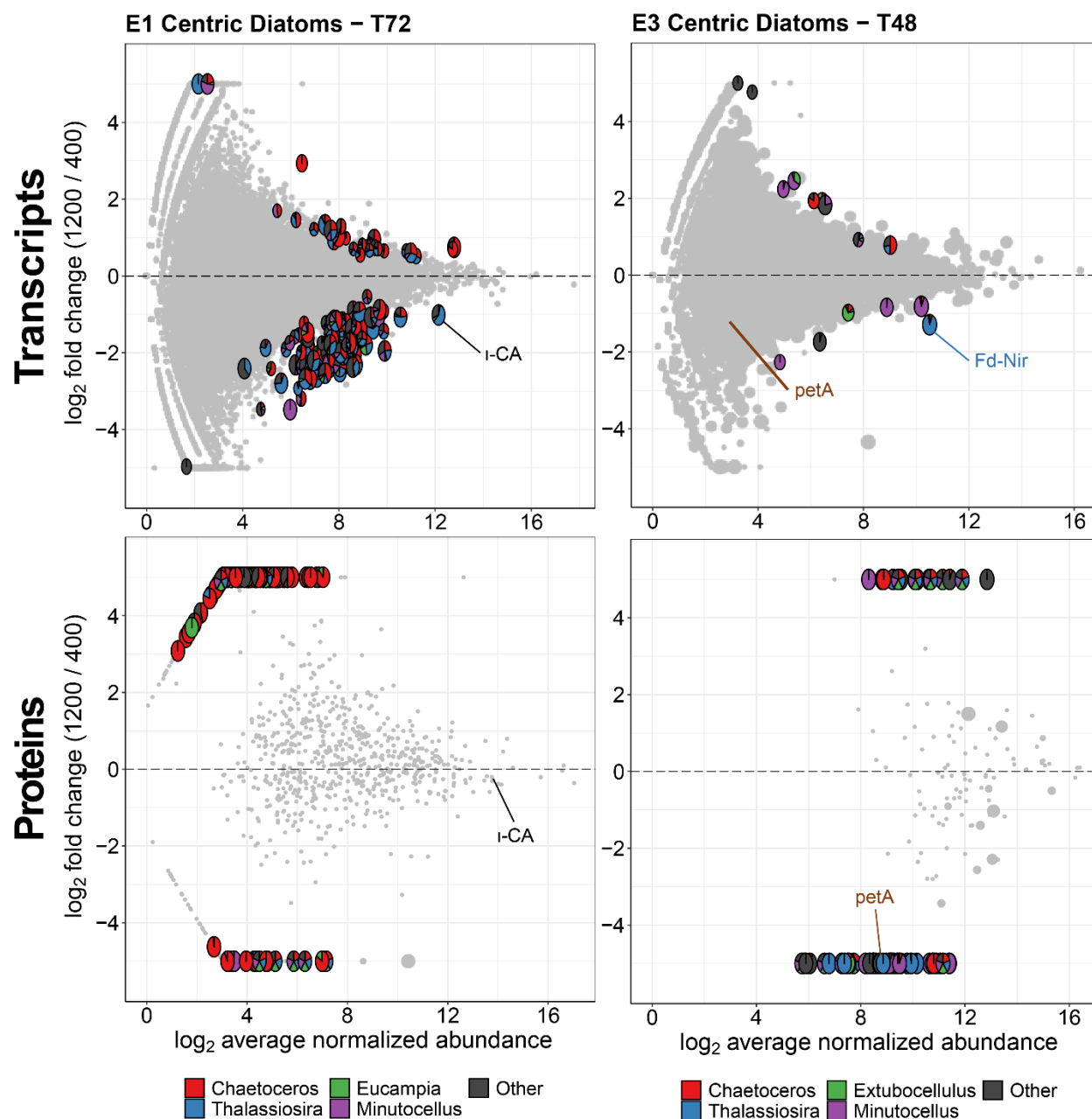
Supplemental Figure 2.20. Maximum likelihood phylogenetic tree as inferred from SLC49 - family amino acid sequences (374 taxa/386 positions) under the LG+C40+F+R10 model as implemented in IQTree 2. The scale bar represents 0.1 amino acid substitutions per position. Branch colours and labels correspond to different taxonomic groups. Numbers above branches indicate ultra-fast Maximum-likelihood bootstrap support. A PDF version of this tree is available in this manuscript's associated figshare repository: doi.org/10.6084/m9.figshare.22589218



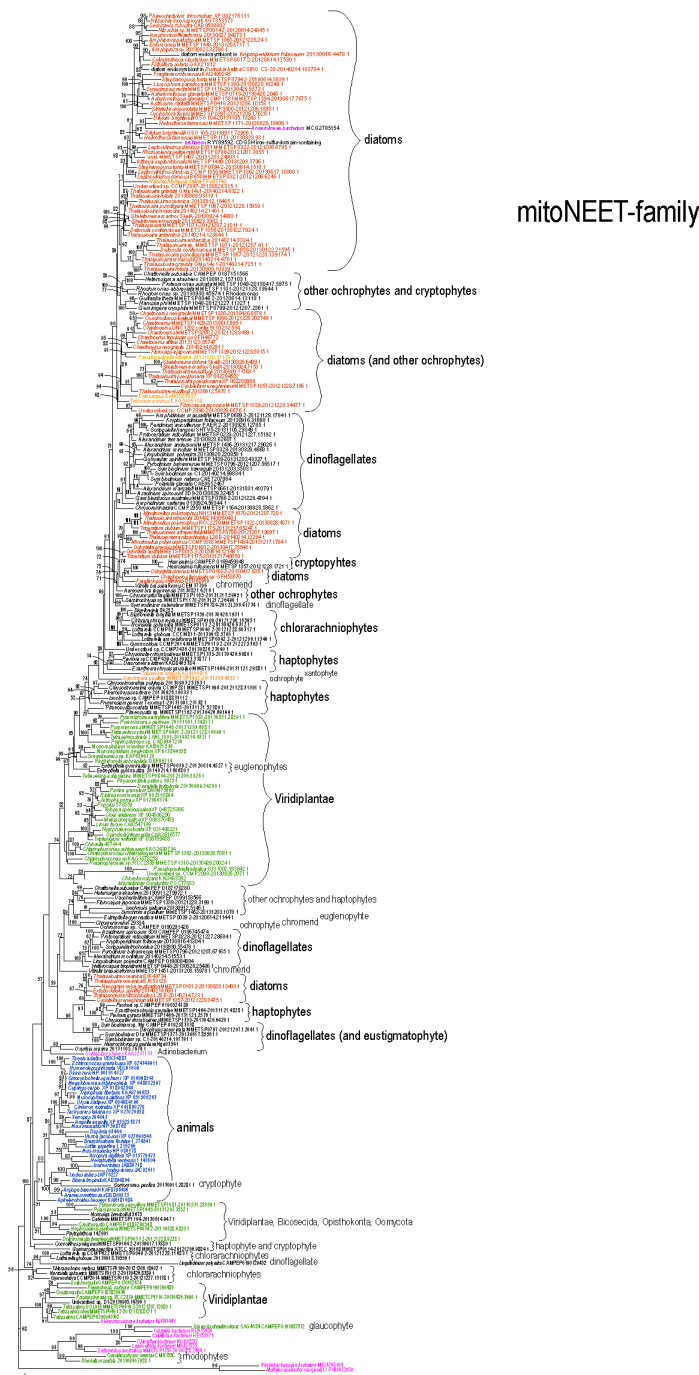
Supplemental Figure 2.21: Normalized transcript expression of ferrichrome binding protein 1 (FBP1) in all diatoms. Error bars represent the standard deviation of the mean ($n = 3$ except the Experiment 3 T0 and Experiment 4 800 ppm samples where $n = 2$).



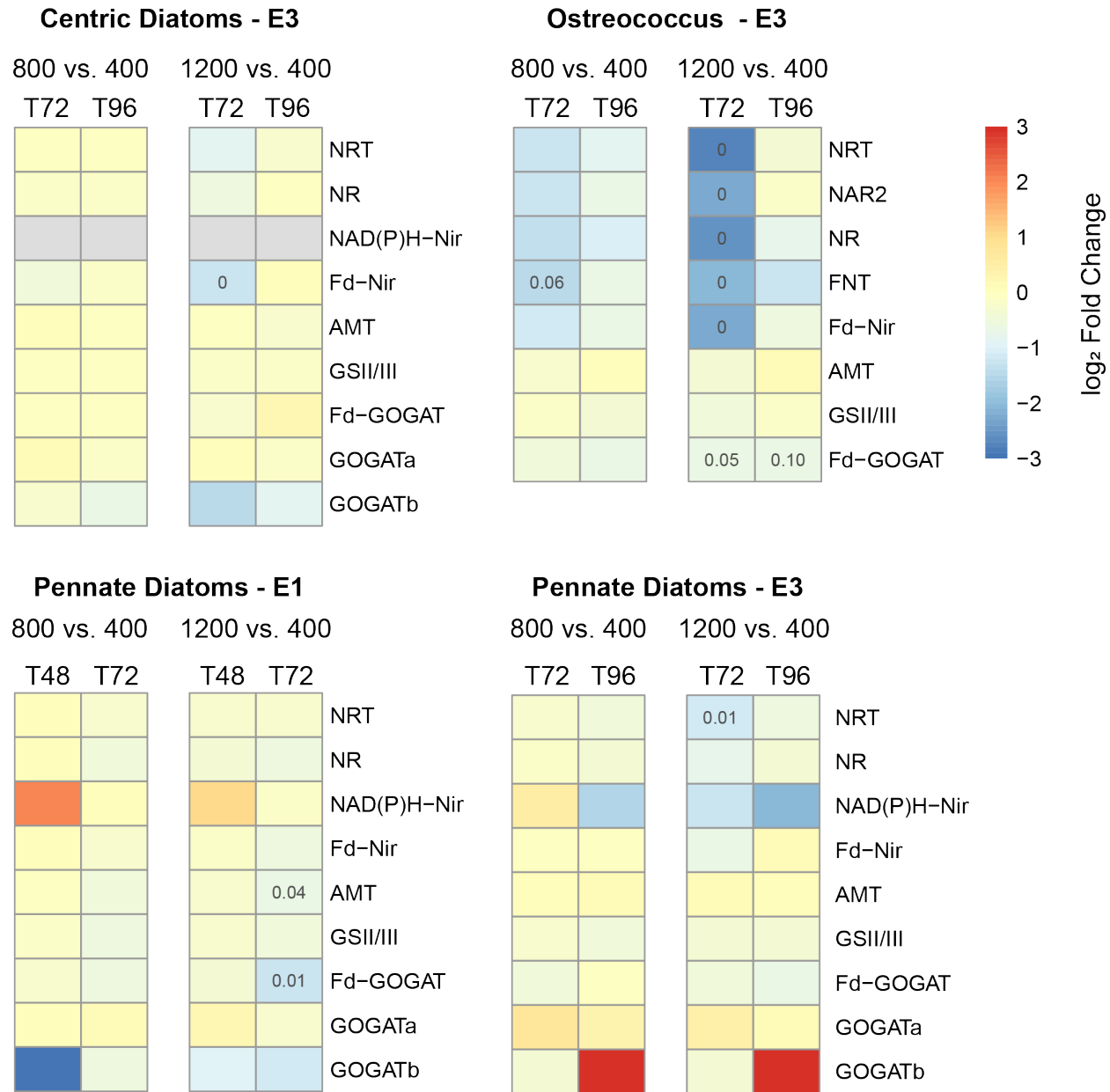
Supplemental Figure 2.22: Normalized (total sum scaled) protein abundances of ferrichrome binding protein 1 (FBP1) in all diatoms at Experiment 1. FBP1 proteins at E1 were either classified as unknown diatoms or centric diatoms with none assigned to pennate diatoms. FBP1 protein expression was not detected at Experiment 3. Error bars represent the standard deviation of the mean ($n = 3$).



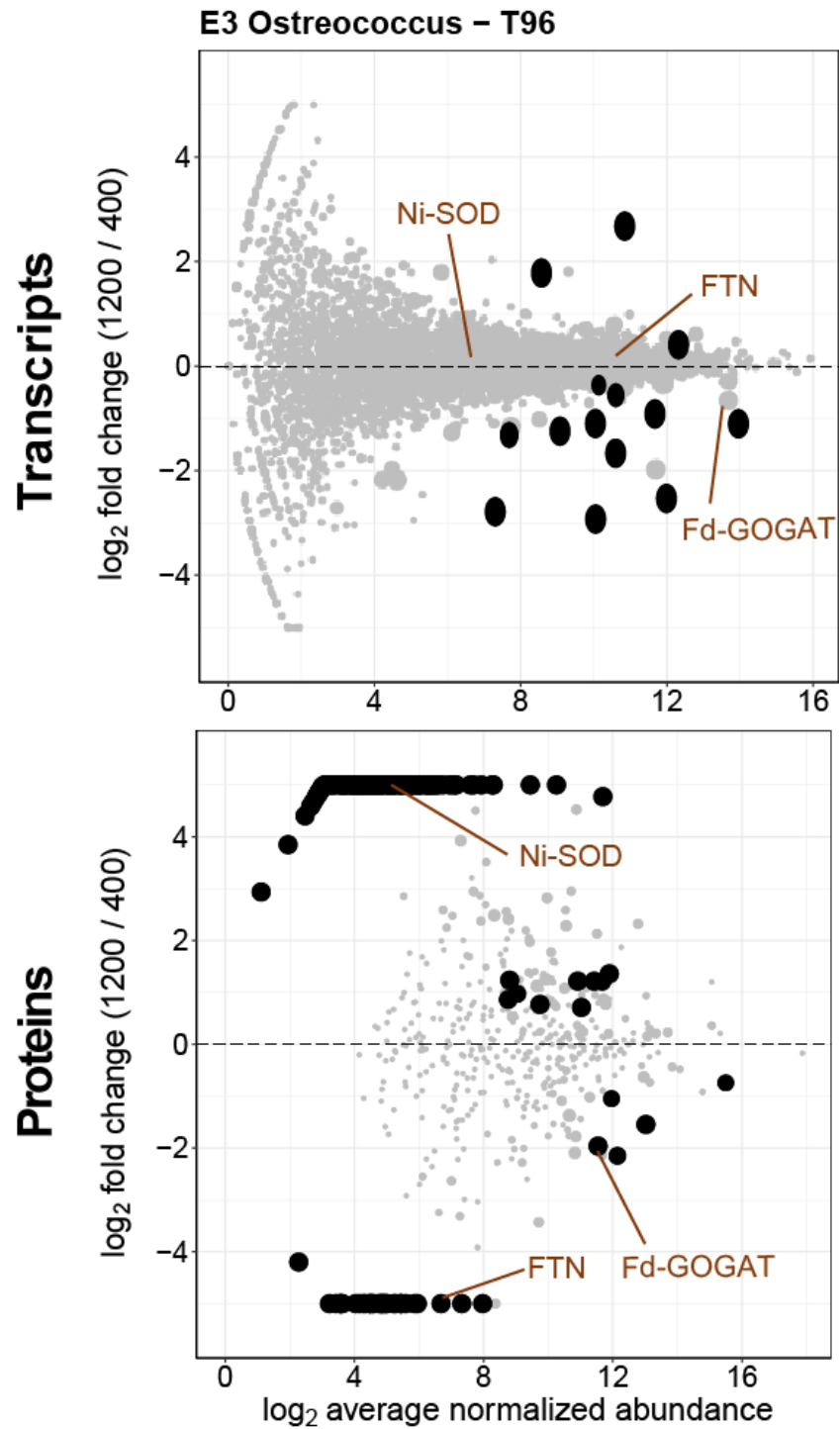
Supplemental Figure 2.23: MA plots for centric diatoms at Experiment 1 (T72) and Experiment 3 (T48). Comparisons are made between the 1200 ppm and 400 ppm treatments for both transcripts (top row) and proteins (bottom row). Differentially expressed genes or proteins ($P < 0.05$) are shown as pies representing the proportion of expression for dominant genera. Pie size is proportional to P -value with larger pies denoting a smaller P -value. Grey points are genes that were not significantly differentially abundance ($P > 0.05$). Data points with $\log_2$ fold changes less than -5 or greater than 5 are plotted at -5 and 5 respectively.



Supplemental Figure 2.24: Maximum likelihood phylogenetic tree as inferred from mitoNEET amino acid sequences (279 taxa/102 positions) under the LG+C40+F+R7 model as implemented in IQTree 2. The scale bar represents 1 amino acid substitutions per position. Branch colours and labels correspond to different taxonomic groups. Numbers above branches indicate ultra-fast Maximum-likelihood bootstrap support. A PDF version of this tree is available in this manuscript's associated figshare repository: doi.org/10.6084/m9.figshare.22589218

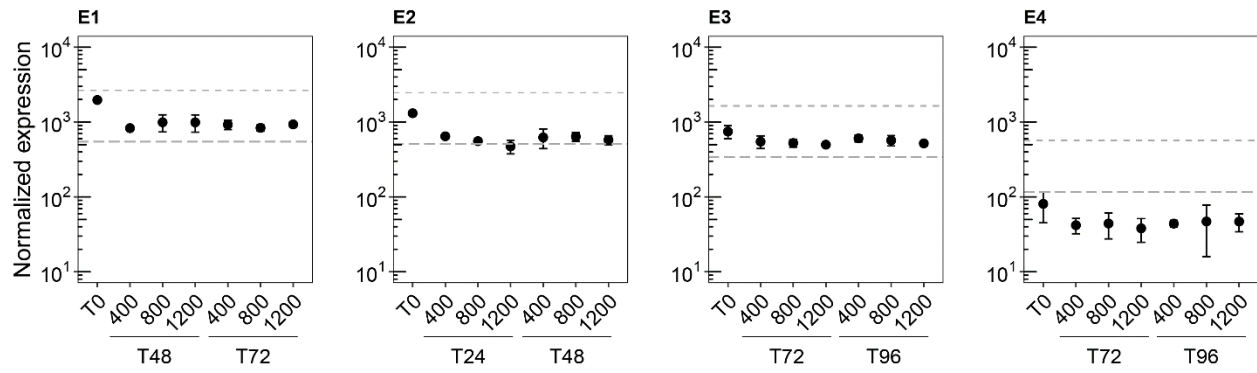


Supplemental Figure 2.25: Differential expression of nitrogen assimilation transcripts for certain taxonomic groups and experiments (E). Colours indicate the log₂ fold change for comparisons listed above each paired time points (T). Adjusted P-values < 0.10 are shown. Genes are abbreviated as follows: NRT nitrate transporter, NAR2 nitrate transport accessory protein, NR nitrate reductase, FNT formate-nitrite transporter, Nir nitrite reductase, AMT ammonium transporter, GSII/III glutamine synthetase GOGAT, glutamate synthase.

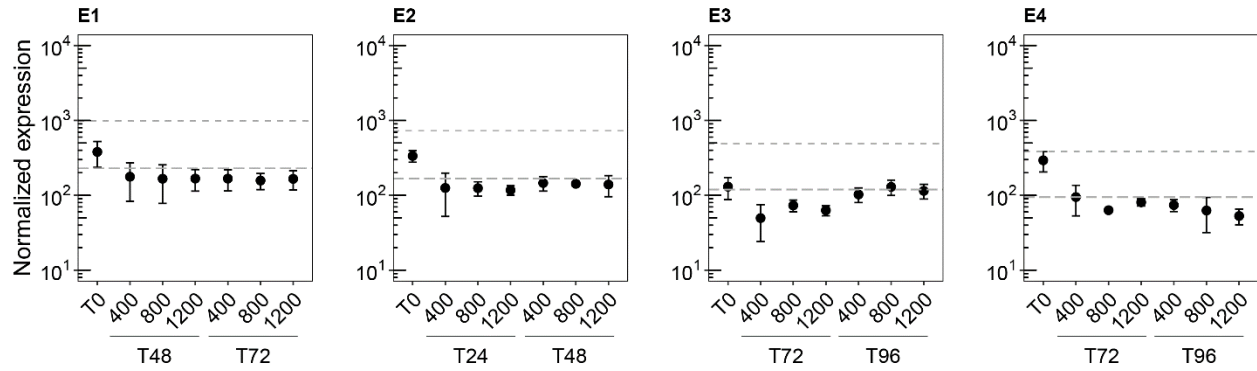


Supplemental Figure 2.26: MA plots for *Ostreococcus* at the second time point (96 hours) at Experiment 3. Black points are shown for significantly differentially expressed transcripts or proteins ($P < 0.05$) and are proportional to the P-value. Data points with log₂ fold changes less than -5 or greater than 5 are plotted at -5 and 5 respectively.

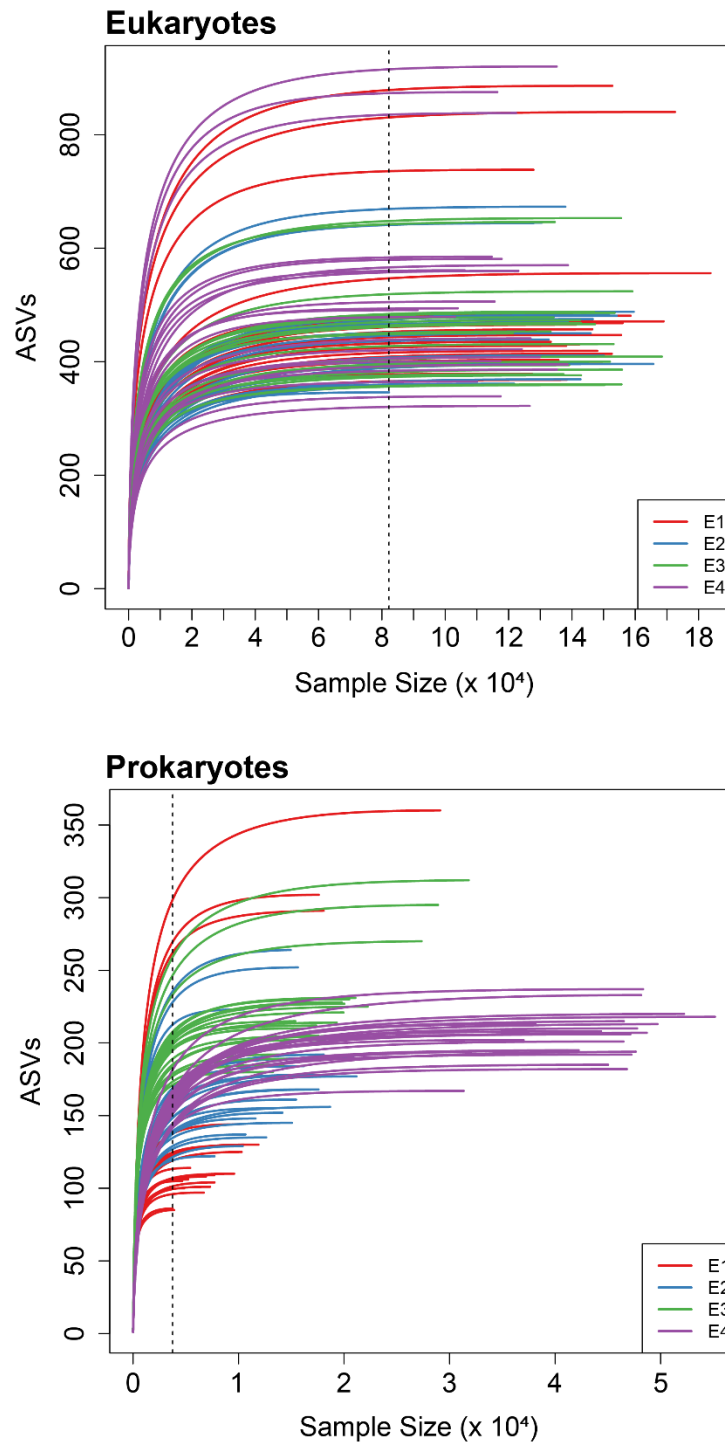
Centric Diatoms



Pennate Diatoms



Supplemental Figure 2.27: SLC4-family bicarbonate transporter transcript expression ($\log_{10}$ scale) in centric diatoms (top) and pennate diatoms (bottom). The horizontal dashed lines indicate the averaged 95th and 99th percentile for transcript abundance in each experiment. Error bars indicate the standard deviation of the mean ($n = 3$ except the Experiment 3 T0 and Experiment 4 800 ppm samples where $n = 2$).



Supplemental Figure 2.28: Rarefaction curves for eukaryotes (18S) and prokaryotes (16S) for each sample. The vertical dashed line indicates the sampling depth for rarefaction. Colours for each sample correspond to the different experiments as shown in the legend.

Supplemental Table 2.1: Reported minimum near-surface pH values in the California Current System. These pH values represent extreme cases and anomalies rather than averages with most variability existing at greater pH values. pCO₂, DIC, and CO₃ ion concentrations were estimated using seacarb(Gattuso et al. 2021) from pH and a constant temperature (11.8°C), salinity (32.7 psu), and total alkalinity (2218 µmol kg⁻¹) as used by Feely et al. (2018).

Data Source	Sampling or Model Years	Depth	pH	pCO₂ (µatm)	DIC (µmol kg⁻¹)	CO₃ (µmol kg⁻¹)
Feely et al. 2008	2007	surface	7.7	944	2162	64
Hauri et al. 2013	2011	surface	7.85	649	2114	88
Feely et al. 2016	2007, 2011-2013		7.7	944	2162	64
Chan et al. 2017	2011-2013	surface	7.43	1818	2241	36
Lilly et al. 2019	2011-2018	16 m	7.64	1094	2180	56

Supplemental Table 2.2: Carbonate system projections for Tier 1 Shared Socioeconomic Pathway (SSP) scenarios representing the wide range of uncertainty by the year 2100 (Gidden et al. 2019). Future atmospheric pCO₂ concentrations (pCO_{2-atm}) were obtained from the SSP Public Database Version 2.0 (Riahi et al. 2017). An 800 ppm atmospheric CO₂ calculation was included as ours and previous experiments have used that concentration as a high CO₂ treatment. For all calculations, constant temperature (11.8°C) and salinity (32.7 psu) as used by Feely et al. (2018) were used. Surface seawater pCO₂ concentrations were calculated with *marelac* v2.1.10 (Soetaert et al. 2020). DIC, pH, and CO₃ ion concentrations were estimated using seacarb (Gattuso et al. 2021) from the estimated pCO_{2-sw} values and a constant total alkalinity of 2234 µmol kg⁻¹, the average value from 2009-2015 CalCOFI samples in the upper 20 m.

SSP	pCO _{2-atm} (ppm)	pCO _{2-sw} (µatm)	DIC (µmol kg ⁻¹)	pH _T	CO ₃ (µmol kg ⁻¹)
1-2.6	432	425	2067	8.017	125
2-4.5	567	557	2108	7.913	101
N/A	800	786	2154	7.776	76
3-7.0	834	820	2160	7.760	73
5-8.5	1089	1070	2192	7.652	58

Supplemental Table 2.3: Average dissolved inorganic carbon (DIC) and total alkalinity (TA) as well as calculated pH, pCO₂, and [CO₃²⁻]. Errors represent the standard deviation of the mean (n = 3).

Exp.	Treat- ment	Hours	DIC ($\mu\text{mol kg}^{-1}$)	TA ($\mu\text{mol kg}^{-1}$)	pH _T	pCO ₂ (μatm)	[CO ₃ ²⁻] ($\mu\text{mol kg}^{-1}$)
E1	T0	0	2123 $\pm$ 2	2247 $\pm$ 2	7.89 $\pm$ 0.00	590 $\pm$ 5	99 $\pm$ 1
	400	48	2048 $\pm$ 20	2245 $\pm$ 12	8.06 $\pm$ 0.03	382 $\pm$ 34	142 $\pm$ 9
		72	2063 $\pm$ 15	2254 $\pm$ 16	8.04 $\pm$ 0.02	396 $\pm$ 24	139 $\pm$ 7
	800	48	2160 $\pm$ 6	2253 $\pm$ 3	7.79 $\pm$ 0.01	763 $\pm$ 22	83 $\pm$ 2
		72	2171 $\pm$ 1	2264 $\pm$ 3	7.79 $\pm$ 0.01	765 $\pm$ 19	84 $\pm$ 2
	1200	48	2199 $\pm$ 8	2230 $\pm$ 6	7.59 $\pm$ 0.02	1241 $\pm$ 71	53 $\pm$ 3
		72	2220 $\pm$ 3	2255 $\pm$ 4	7.61 $\pm$ 0.01	1200 $\pm$ 32	56 $\pm$ 1
E2	T0	0	2088 $\pm$ 8	2242 $\pm$ 10	7.94 $\pm$ 0.04	516 $\pm$ 52	117 $\pm$ 11
	400	24	2082 $\pm$ 3	2262 $\pm$ 4	8.02 $\pm$ 0.00	428 $\pm$ 3	132 $\pm$ 1
		48	2086 $\pm$ 7	2266 $\pm$ 5	8.02 $\pm$ 0.01	427 $\pm$ 8	132 $\pm$ 1
	800	24	2180 $\pm$ 5	2264 $\pm$ 4	7.76 $\pm$ 0.02	832 $\pm$ 49	78 $\pm$ 4
		48	2181 $\pm$ 2	2265 $\pm$ 4	7.76 $\pm$ 0.01	830 $\pm$ 28	78 $\pm$ 2
	1200	24	2227 $\pm$ 3	2265 $\pm$ 3	7.61 $\pm$ 0.01	1192 $\pm$ 39	57 $\pm$ 2
		48	2230 $\pm$ 2	2268 $\pm$ 2	7.61 $\pm$ 0.01	1186 $\pm$ 36	58 $\pm$ 2
E3	T0	0	2022 $\pm$ 4	2244 $\pm$ 2	8.06 $\pm$ 0.01	377 $\pm$ 5	159 $\pm$ 1
	400	48	2079 $\pm$ 3	2259 $\pm$ 3	8.02 $\pm$ 0.01	425 $\pm$ 5	132 $\pm$ 1
		96	2085 $\pm$ 2	2272 $\pm$ 6	8.03 $\pm$ 0.01	410 $\pm$ 8	137 $\pm$ 3
	800	48	2176 $\pm$ 2	2261 $\pm$ 6	7.76 $\pm$ 0.01	816 $\pm$ 26	79 $\pm$ 3
		96	2181 $\pm$ 2	2274 $\pm$ 2	7.79 $\pm$ 0.00	777 $\pm$ 9	83 $\pm$ 1
	1200	48	2227 $\pm$ 1	2262 $\pm$ 4	7.61 $\pm$ 0.01	1208 $\pm$ 31	56 $\pm$ 1
		96	2230 $\pm$ 3	2271 $\pm$ 4	7.62 $\pm$ 0.02	1160 $\pm$ 53	59 $\pm$ 2
E4	T0	0	2016 $\pm$ 3	2221 $\pm$ 2	8.06 $\pm$ 0.01	374 $\pm$ 6	148 $\pm$ 2
	400	48	2062 $\pm$ 1	2242 $\pm$ 4	8.02 $\pm$ 0.01	418 $\pm$ 9	131 $\pm$ 2
		96	2067 $\pm$ 3	2245 $\pm$ 4	8.02 $\pm$ 0.00	424 $\pm$ 2	130 $\pm$ 1
	800	48	2157 $\pm$ 4	2241 $\pm$ 7	7.77 $\pm$ 0.01	807 $\pm$ 13	77 $\pm$ 1
		96	2162 $\pm$ 5	2245 $\pm$ 4	7.76 $\pm$ 0.00	813 $\pm$ 9	77 $\pm$ 1
	1200	48	2207 $\pm$ 2	2241 $\pm$ 3	7.60 $\pm$ 0.00	1213 $\pm$ 5	54 $\pm$ 0
		96	2209 $\pm$ 9	2247 $\pm$ 4	7.61 $\pm$ 0.02	1179 $\pm$ 55	56 $\pm$ 2

Supplemental Table 2.4: Comparison between the number of significantly differentially expressed transcript (T) or protein (P) clusters for dominant taxonomic groups at each experiment (DESeq2 Wald Test, Benjamini & Hochberg adjusted $P < 0.05$).

Group	Exp. & Time-point	Total transcripts	Total proteins	400 ppm transcripts	400 ppm proteins	400 ppm both	1200 ppm transcripts	1200 ppm proteins	1200 ppm both
Centric Diatoms	E1-T1	27,446	926	40	57	0	155	46	1
	E1-T2	22,834	639	105	18	1	29	115	1
	E3-T1	13,603	160	6	46	0	9	15	0
	E3-T2	14,489	368	1	27	0	30	84	4
Pennate diatoms	E1-T1	19,518	267	0	26	0	1	26	0
	E1-T2	26,845	995	8	24	0	3	12	0
	E3-T1	12,722	89	2	27	0	0	15	0
	E3-T2	13,236	195	0	25	0	2	43	0
Ostreococcus	E3-T1	5,670	594	27	95	5	5	27	1
	E3-T2	5,696	595	11	40	0	3	145	0
Pelagomonas	E3-T1	7,671	83	0	16	0	0	11	0
	E3-T2	7,564	61	0	5	0	0	30	0

Supplemental Table 2.5: The number of significantly differentially expressed transcript (T) or protein (P) clusters for dominant taxonomic groups at each experiment (DESeq2 Wald Test, Benjamini & Hochberg adjusted $P < 0.05$). These data are presented as a heatmap in Fig. 2.4a.

		800 v 400 ppm				1200 v 400 ppm			
		T1		T2		T1		T2	
Group	Exp.	400	800	400	800	400	1200	400	1200
Centric Diatoms	E1 (T)	1	0	2	3	40	155	105	29
	E1 (P)	-	-	-	-	57	46	18	115
	E2 (T)	1	7	0	0	0	0	0	0
	E3 (T)	1	0	0	3	6	9	1	30
	E3 (P)	-	-	-	-	46	15	27	84
	E4 (T)	14	1	0	0	3	10	1	1
Pennate Diatoms	E1 (T)	0	1	0	0	0	1	8	3
	E1 (P)	-	-	-	-	26	26	24	12
	E2 (T)	0	1	0	0	1	4	0	1
	E3 (T)	0	0	0	1	2	0	0	2
	E3 (P)	-	-	-	-	27	15	25	43
	E4 (T)	0	0	0	0	0	4	0	1
Ostreococcus	E3 (T)	15	4	10	1	27	5	11	3
	E3 (P)	-	-	-	-	95	27	40	145
Pelagomonas	E3 (T)	0	0	0	0	0	0	0	0
	E3 (P)	-	-	-	-	16	11	5	30
	E4 (T)	71	25	0	0	0	0	0	0

Supplemental Table 2.6: Liquid chromatography flow rates and solvent composition settings.
Arrows indicate linear increase or decreases in Solvent B* and A** composition.

Time (minutes)	Flow (µl/min)	% Solvent B	% Solvent A
0-15	0.3	5	95
15.1 - 90	0.25	5 → 30	95 → 70
90.1 - 102	0.25	30 → 55	70 → 45
102.1 - 106	0.3	55 → 95	45 → 5
106.1 - 110	0.3	95	5
111 - 125	0.3	5	95

*Solvent A: H₂O, 0.1% Formic Acid

** Solvent B: Acetonitrile, 0.1% Formic Acid

Supplemental Table 2.7: Mass spectrometer settings.

Parameter	Setting
TopN	8
Intensity Threshold	8.3e4
Dynamic Exclusion	30s
MS1 Scan Resolution	140,000
MS1 Scan Range	400 to 2000 m/z
MS1 Automatic Gain Control Target	3e6
MS2 Scan Resolution	17,500
MS2 Scan Range	200 to 2000 m/z
M2 Automatic Gain Control Target	1e6
MS2 Isolation Window	2.0 m/z

Supplemental References

Chan, F., J. A. Barth, C. A. Blanchette, R. H. Byrne, F. Chavez, O. Cheriton, R. A. Feely, G. Friederich, B. Gaylord, T. Gouhier, S. Hacker, T. Hill, G. Hofmann, M. A. McManus, B. A. Menge, K. J. Nielsen, A. Russell, E. Sanford, J. Sevadjan, and L. Washburn. 2017. Persistent spatial structuring of coastal ocean acidification in the California Current System. *Sci Rep* **7**: 2526.

Feely, R. A., S. R. Alin, B. Carter, N. Bednaršek, B. Hales, F. Chan, T. M. Hill, B. Gaylord, E. Sanford, R. H. Byrne, C. L. Sabine, D. Greeley, and L. Juranek. 2016. Chemical and biological impacts of ocean acidification along the west coast of North America. *Estuar Coast Shelf Sci* **183**: 260-270.

Feely, R. A., R. R. Okazaki, W.-J. Cai, N. Bednaršek, S. R. Alin, R. H. Byrne, and A. Fassbender. 2018. The combined effects of acidification and hypoxia on pH and aragonite saturation in the coastal waters of the California current ecosystem and the northern Gulf of Mexico. *Cont Shelf Res* **152**: 50-60.

Feely, R. A., C. L. Sabine, J. M. Hernandez-Ayon, D. Ianson, and B. Hales. 2008. Evidence for Upwelling of Corrosive "Acidified" Water onto the Continental Shelf. *Science* **320**: 1490-1492.

Gattuso, J.-P., J.-M. Epitalon, H. Lavigne, J. Orr, B. Gentili, M. Hagens, A. Hofmann, J.-D. Mueller, A. Proye, and J. Rae. 2021. Package 'seacarb'.

Gidden, M. J., K. Riahi, S. J. Smith, S. Fujimori, G. Luderer, E. Kriegler, D. P. van Vuuren, M. van den Berg, L. Feng, D. Klein, K. Calvin, J. C. Doelman, S. Frank, O. Fricko, M. Harmsen, T. Hasegawa, P. Havlik, J. Hilaire, R. Hoesly, J. Horing, A. Popp, E. Stehfest, and K. Takahashi. 2019. Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century. *Geosci Model Dev* **12**: 1443-1475.

Hauri, C., N. Gruber, M. Vogt, S. C. Doney, R. A. Feely, Z. Lachkar, A. Leinweber, A. M. P. McDonnell, M. Munnich, and G. K. Plattner. 2013. Spatiotemporal variability and long-term trends of ocean acidification in the California Current System. *Biogeosciences* **10**: 193-216.

Lilly, L. E., U. Send, M. Lankhorst, T. R. Martz, R. A. Feely, A. J. Sutton, and M. D. Ohman. 2019. Biogeochemical Anomalies at Two Southern California Current System Moorings During the 2014–2016 Warm Anomaly-El Niño Sequence. *J Geophys Res Oceans* **124**: 6886-6903.

Riahi, K., D. P. van Vuuren, E. Kriegler, J. Edmonds, B. C. O'Neill, S. Fujimori, N. Bauer, K. Calvin, R. Dellink, O. Fricko, W. Lutz, A. Popp, J. C. Cuaresma, S. Kc, M. Leimbach, L. Jiang, T. Kram, S. Rao, J. Emmerling, K. Ebi, T. Hasegawa, P. Havlik, F. Humpenöder, L. A. Da Silva, S. Smith, E. Stehfest, V. Bosetti, J. Eom, D. Gernaat, T. Masui, J. Rogelj, J. Stremler, L. Drouet, V. Krey, G. Luderer, M. Harmsen, K. Takahashi, L. Baumstark, J. C. Doelman, M. Kainuma, Z. Klimont, G. Marangoni, H. Lotze-Campen, M. Obersteiner, A. Tabeau, and M. Tavoni. 2017. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Glob Environ Change* **42**: 153-168.

Soetaert, K., T. Petzoldt, and F. Meysman. 2020. marelac: Tools for aquatic sciences. R package version 2.1.10.

CHAPTER 3 - RELATIONSHIPS BETWEEN PHYTOPLANKTON PIGMENTS AND DNA- OR RNA-BASED ABUNDANCES SUPPORT ECOLOGICAL APPLICATIONS

Robert H. Lampe, Ariel J. Rabines, Steffaney M. Wood, Anne Schulberg, Ralf Goericke, Pratap Venepally, Hong Zheng, Michael R. Stukel, Michael R. Landry, Andrew D. Barton, Andrew E. Allen

Abstract

Observations of phytoplankton abundances and community structure are critical towards understanding marine ecosystems. Common approaches to determine group-specific abundances include measuring phytoplankton pigments via high-performance liquid chromatography and DNA-based metabarcoding. Increasingly, mRNA abundances via metatranscriptomics are also employed. As phytoplankton pigments are used to develop and validate remote sensing algorithms, further comparisons between pigments and other metrics are needed to validate the extent to which these measurements agree for group-specific abundances; however, most previous comparisons have been hindered by metabarcoding and metatranscriptomics solely producing relative abundance data. By employing quantitative approaches that express both 18S rDNA and total mRNA as concentrations, we show that these measurements are related for several eukaryotic phytoplankton groups. We further propose that integration of these can be used to examine ecological patterns more deeply. For example, productivity-diversity relationships of both the whole community and individual groups show a dinoflagellate-driven negative trend rather than the commonly-found unimodal pattern. Pigments are also shown to relate to certain harmful algal bloom-forming taxa as well as the expression of sets of genes. Altogether, these results suggest that potential models of pigment concentrations via hyperspectral remote sensing may enable improved assessments of global phytoplankton community structure, including the detection of harmful algal blooms, and support the development of ecosystem models.

1 Introduction

Marine phytoplankton comprise both prokaryotic cyanobacteria and diverse lineages of eukaryotes with distinct evolutionary histories (Pierella Karlusich et al. 2020). Collectively, they are responsible for approximately 50% of global primary production, support marine food webs and extensively contribute to biogeochemical cycling (Huang et al. 2021). Their distinct evolutionary trajectories have also given rise to different traits, functional roles, and niches. For example, diatoms (Bacillariophyta) tend to dominate under well-mixed, nutrient-rich conditions, are considered principal contributors to carbon export with fast sinking rates, and are a significant component of the silicon cycle due to their formation of silica cell walls (Agusti et al. 2015; Bowler et al. 2010; Guidi et al. 2016). Dinoflagellates (Dinoflagellata) may be major contributors to carbon export in oligotrophic regions, and many species are mixotrophs, which enables them to serve multiple important roles in marine food webs (Guidi et al. 2016; Stoecker et al. 2017). Prymnesiophytes (Haptophyta) include coccolithophores with calcium carbonate shells, the formation of which influences upper ocean alkalinity and air-sea carbon dioxide exchange (Rost and Riebesell 2004). The prymnesiophytes also include the bloom-forming *Phaeocystis* spp. which are a major source of the sulfur compound, dimethylsulfoniopropanate (DMSP), that is hypothesized to influence cloud formation and climate (Smith Jr and Trimborn 2023). Cryptophytes (Cryptophyta) are considered to be ubiquitous and may be highly abundant in the Sargasso Sea and Southern Ocean (Cotti-Rausch et al. 2016; Mendes et al. 2018; Richardson 2022). Lastly, the chlorophytes (Chlorophyta) and pelagophytes (Pelagophyceae) are additional picoeukaryotes that are important primary producers, particularly in the open ocean (Guérin et al. 2022; Lopes dos Santos et al. 2017; Not et al. 2012; Worden et al. 2012).

Due to these contributions to ecosystem function and differing roles, assessing phytoplankton abundances and disentangling phytoplankton community composition (PCC) is necessary to understand marine environments (Cetinić et al. 2024). For example, current Earth systems models (ESMs) are unable to confidently project if primary productivity will increase or decrease under future climate scenarios (Kwiatkowski et al. 2020). By examining the effects of natural climate variability and anthropogenic change on phytoplankton communities as well as their physiology, more accurate and detailed representation of different groups can be generated to improve these model predictions (Cetinić et al. 2024; Tagliabue 2023). Knowledge of PCC may also further help estimate carbon export flux (Kramer et al. 2024b). Additionally, monitoring for harmful algal blooms (HABs) is important for human health as well as understanding their impacts on fisheries and ecosystems (Anderson et al. 2012), and fisheries managers may benefit from knowledge of PCC to aid stock assessments and identify suitable fishing zones (Sathyendranath et al. 2023; Satterthwaite et al. 2023).

A common approach to estimate phytoplankton abundances and determine PCC is the measurement of phytoplankton pigments via high-performance liquid chromatography (HPLC). Specifically, certain accessory pigments can be used as proxies, or diagnostic pigments, to determine the abundances of specific groups, even though some of these pigments are shared among groups (Jeffrey et al. 2011; Kramer and Siegel 2019). As pigments directly alter the shapes and magnitudes of remote sensing reflectance spectra, HPLC-measured pigments are also extensively used to develop and validate satellite-based remote sensing algorithms. Given the recent availability of hyperspectral remote sensing reflectance data from NASA's Plankton, Aerosol, Cloud, ocean Ecology (PACE) mission, the continued development of remote sensing

algorithms based on HPLC pigments may enable global scale estimation of phytoplankton pigment concentrations and PCC (Cetinić et al. 2024; Kramer et al. 2022).

As a result, it is important to compare phytoplankton pigments to other metrics of PCC. If other metrics agree with pigments, they may also be useful for validating remote sensing algorithms. Conversely, it is important to compare other metrics of PCC to phytoplankton pigments as it remains uncertain whether indirect approaches such as DNA and RNA sequencing are reflective of phytoplankton biomass. DNA metabarcoding is increasingly being used to determine PCC, primarily by sequencing hypervariable regions of the 16S gene for prokaryotes and phytoplankton plastids or the 18S gene for eukaryotes (Lopes dos Santos et al. 2022), and it is increasingly common to use environmental mRNA sequencing, or metatranscriptomics, to determine PCC (Cohen et al. 2022).

Initial comparisons between phytoplankton pigments and DNA indicate that the metrics are correlated in some cases; however, these comparisons have been hindered by the compositional nature of sequencing data (Catlett et al. 2023; Kramer et al. 2024a). By default, DNA and RNA sequencing produce relative abundance data, which complicates their interpretability and can lead to spurious correlations (Gloor et al. 2017). This is because there are multiple scenarios that could lead to a taxonomic group having the same or different relative abundances among samples, and potential differences in relative abundances are not necessarily reflective of differences in their true abundances. For example, one group's true abundance could be the same in two samples while another group's is lower in just one sample leading to a higher relative abundance of the first group even though the two abundances are equal. Alternatively, both groups' true abundances could be higher in one sample but to different degrees, resulting in lower relative abundances in one group even though its true abundance has increased.

To compare phytoplankton pigments and relative DNA abundances, previous studies have also normalized diagnostic pigment concentrations to total chlorophyll a concentrations or the sum of diagnostic pigments, thereby making both DNA and pigment data relative quantities (Catlett et al. 2023; Kramer et al. 2024a). However, the use of these approaches while only using relative abundances of prokaryotic or eukaryotic taxa presents other issues. As an example, high total chlorophyll a could be driven by relatively high cyanobacterial abundances that would not be captured in 18S sequencing, which targets eukaryotes. If the eukaryotic phytoplankton community is then dominated by one group, this group would have a high relative abundance that is coupled to low concentrations of their diagnostic pigment relative to total chlorophyll a. The use of 16S alone to capture both prokaryotic cyanobacteria and eukaryotic plastids may circumvent this issue, but references for plastid 16S sequences are not as comprehensive and are influenced by potentially greater variability in plastid copy number (Decelle et al. 2015; Lopes dos Santos et al. 2022). Alternatively, simultaneous metabarcoding of the 16S and 18S rRNA gene with three domain primers addresses this issue, but it may also require greater sequencing depth to capture the diversity of the eukaryotic community (Yeh et al. 2021). Even if eukaryotic relative abundances are only compared to the sum of eukaryotic diagnostic pigments, the DNA-based relative abundances are still influenced by non-photosynthetic taxa. Altogether, these problems with normalization and potential discrepancies between true abundances and relative metrics can be severe issues that produce misleading correlations.

The estimation of absolute abundances via internal standards avoids these issues and enables comparisons between gene or transcript concentrations and phytoplankton pigment concentrations (Cohen et al. 2022; Lin et al. 2019). This approach was used in a study in the Western Antarctic Peninsula to compare 18S DNA concentrations to group-specific chlorophyll

a concentrations (Lin et al. 2019). With a small sample size ($n = 16$) and cryptophyte-dominated communities, strong relationships for cryptophytes and *Phaeocystis* were observed, but not with diatoms. Other groups were not examined, and it remains unclear if these relationships persist with larger sample sizes, other regions with different communities, or across the entire spectrum of phytoplankton abundances. Furthermore, it remains unclear if absolute quantities of mRNA from metatranscriptomics and phytoplankton pigments relate to one another.

Here we investigate relationships between phytoplankton pigments and quantitative DNA- or RNA-based abundances with paired samples collected over a seven-year period on seasonal California Cooperative Oceanic Fisheries Investigations (CalCOFI) surveys. The CalCOFI sampling area is within the California Current Ecosystem, a coastal upwelling biome associated with an eastern boundary current. Although the sampling area is largely restricted to the southern California Current (Figs. 3.1A and 3.1B), it is along a major biogeographic boundary at Point Conception, and contains phytoplankton taxa commonly associated with both the central North Pacific Subtropical Gyre and the subarctic Northeast Pacific; thus, it captures communities associated with the broader region (Checkley and Barth 2009; Venrick 1998). The range in phytoplankton abundances closely aligns with those observed globally and is captured by sampling the gradients from nearshore stations influenced by upwelling to offshore oligotrophic stations (Deutsch et al. 2021; James et al. 2022; Venrick 2002). CalCOFI additionally conducts four cruises per year, capturing differences from seasonal and interannual variability. Beyond examining PCC, we explore additional applications stemming from relationships between phytoplankton pigments and DNA- or RNA-based abundances including investigating productivity-diversity relationships, forecasting harmful algal blooms, and inferring expression of specific genes.

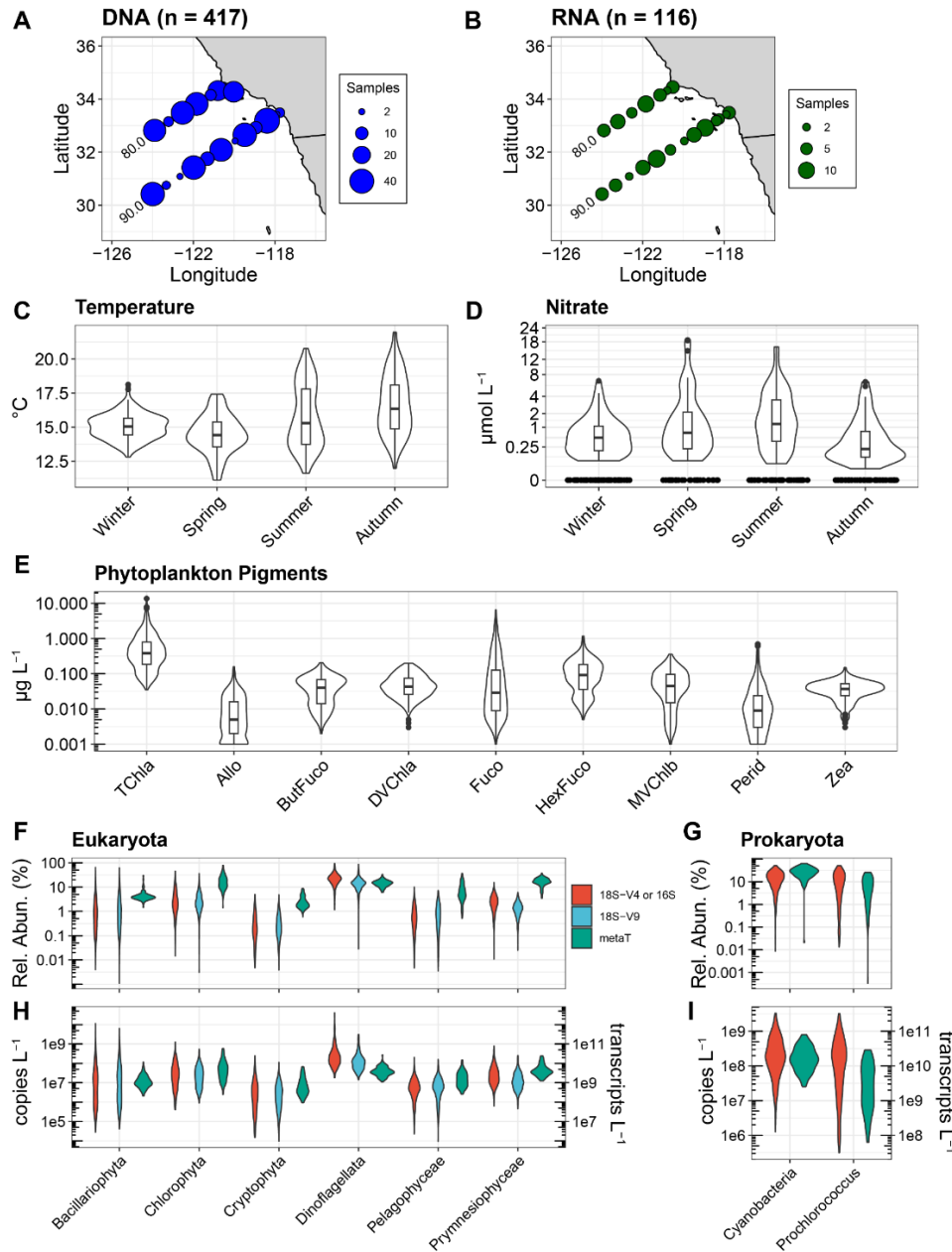


Figure 3.1: Overview of CalCOFI samples used in this study. (A, B) Locations and quantity of samples for DNA and RNA with corresponding phytoplankton pigment samples. (C, D) Temperature (°C) and nitrate concentrations (μmol L⁻¹) across seasons for the paired DNA and pigment samples used here. Corresponding data for the RNA samples are shown in Supplemental Fig. 3.1. (E) Phytoplankton pigment concentrations. Pigment abbreviations: TChla, total chlorophyll a; Allo, alloxanthin; ButFuco, 19'-butanoyloxyfucoxanthin; DVChla, divinyl chlorophyll a; Fuco, fucoxanthin; HexFuco; 19'-hexanoyloxyfucoxanthin; MVChlb, monovinyl chlorophyll b; Perid, peridinin; Zea, zeaxanthin. (F, G) Relative abundances of different phytoplankton groups using the 18S-V4 (red), 18S-V9 (blue), or transcript abundance (metaT, green). For prokaryota, the 16S V4-V5 region was used and is shown in red. (H, I) Absolute abundances for 18S or 16S rRNA genes (copies L⁻¹, left y-axis) and transcripts (transcripts L⁻¹, right y-axis). Different metrics are colored as in panels F and G.

2 Materials and Methods

2.1 Sample collection and biogeochemical measurements

From select stations on seasonal CalCOFI cruises from 2014 to 2020, 417 DNA and 118 RNA samples were collected concurrently with phytoplankton pigments within the euphotic zone along CalCOFI lines 80.0 and 90.0 and at the Santa Barbara Basin (Station 81.8 46.9) (Figs. 3.1A and 3.1B). These data represent only a subset of the on-going NOAA-CalCOFI Ocean Genomics (NCOG) time series, with no DNA samples from 2017 and only RNA samples from 2017-2020 to examine samples only where quantitative approaches for DNA and RNA were employed with concurrent sampling of phytoplankton pigments (James et al. 2022). CalCOFI cruises survey the California Current Ecosystem in a standardized grid pattern with most stations in the southern California Current region. During this time period, winter cruises occurred during January and February, spring cruises occurred during April, summer cruises occurred in July and August, and autumn cruises occurred in October and November. There was no Spring cruise during 2020 due to the COVID-19 pandemic. At each station, seawater was collected from the near-surface (normally 10 m) and the subsurface chlorophyll maximum layer (SCML) with a CTD rosette for pigments, DNA, RNA, and flow cytometry. Temperature data are derived from duplicate Seabird SBE 3Plus sensors on a Seabird 911+ CTD. Macronutrient concentrations were measured on a QuAatro continuous segmented flow autoanalyzer (SEAL Analytical) alongside reference materials (KANSO technos) (Armstrong et al. 1967; Gordon et al. 1992). The nitracline depth is defined as the depth where nitrate first exceeds 1 μM .

Primary productivity was estimated via ^{14}C uptake at select stations. Briefly, seawater was collected from six depths representing 56%, 30%, 10%, 3%, 1%, and 0.3% surface light levels shortly before local apparent noon and dispensed into triplicate 250 mL polycarbonate

bottles (two light bottles and one dark control). Bottles were inoculated with $\text{NaH}^{14}\text{CO}_3$ and incubated until civil twilight in tubes with flow-through seawater and neutral-density screens to simulate in situ light levels. Following incubation, samples were filtered onto $0.45\ \mu\text{m}$ HA filters (Millipore), acidified with HCl, immersed in scintillation fluor, and measured with a scintillation counter once back onshore. Half-day productivity at each depth was averaged between the two light bottles and corrected with the dark uptake bottle then multiplied by 1.8 to obtain 24 hour productivity (Eppley 1992). When comparing productivity to diversity from DNA, samples from the entire NCOG dataset (2014 to 2020) that were closest to the collection depths of productivity samples were used ($n = 757$).

2.2 High-performance liquid chromatography pigment analysis

Phytoplankton pigment concentrations were determined with high-performance liquid chromatography (HPLC). Samples were collected with 0.5, 1.04, or 2.2 L opaque brown bottles depending on the fluorescence measured by the CTD and filtered onto 25 mm GF/F filters under low vacuum pressure ($\leq 40\ \text{mm Hg}$). Once completed filtering, the filters were carefully folded in half, blotted on a paper towel to remove excess water, and stored in 2 mL cryovials in liquid nitrogen until analysis at the Horn Point Analytical Services Laboratory at the University of Maryland. Taxon-specific contributions to total chlorophyll a concentrations were determined with phytoclass v1.0.0 (Hayward et al. 2023).

2.3 Nucleic acid sample collection and extraction

For DNA samples, 0.2 to 10.4 L (mean = 3.3 L) of seawater was filtered onto $0.22\ \mu\text{m}$ Sterivex™ filters. RNA samples were collected simultaneously following the same approach but

with generally higher volumes (2.0 to 4.8 L, mean = 4.0 L). RNA samples were also only collected near local apparent noon to minimize bias from the diel cycle.

Following filtration, samples were immediately flash frozen in liquid nitrogen and stored at -80°C. DNA was extracted with the Macherey-Nagel NucleoMag Plant kit on an Eppendorf epMotion 5075TMX and assessed on a 1.8% agarose gel. At the start of DNA extraction during the addition of lysis buffer, 1.74 to 3.78 ng of *Schizosaccharomyces pombe* genomic DNA was added to each sample as an internal standard (Lin et al. 2019). RNA was also extracted on the Eppendorf epMotion but with the Machery-Nagel NucleoMag RNA kit. As an internal standard, 2 to 5 billion copies of Invitrogen™ ArrayControl™ RNA Spikes #1 and #8 were added to the lysis buffer of each sample at a 2.66:1 ratio.

2.4 Amplicon library preparation and sequencing

Amplicon libraries separately targeting the V4-V5 region of the 16S rRNA gene, the V4 and V9 regions of the 18S rRNA gene, and the ITS2 gene from the diatom genus *Pseudo-nitzschia* were constructed via a one-step PCR with the TruFi DNA Polymerase PCR kit to simultaneously amplify the region of interest and barcode each sample. For 16S, the 515F-Y (5'-GTG YCA GCM GCC GCG GTA A-3') and 926R (5'-CCG YCA ATT YMT TTR AGT TT-3') primer set was used (Parada et al. 2016). For 18S-V4, the V4F (5'-CCA GCA SCY GCG GTA ATT CC-3') and V4RB (5'-CCA GCA SCY GCG GTA ATT CC-3') primer set was used (Berdjeb et al. 2018). For 18S-V9, the 1389F (5'-TTG TAC ACA CCG CCC-3') and 1510R (5'-CCT TCY GCA GGT TCA CCT AC-5') primer set was used (Amaral-Zettler et al. 2009). For *Pseudo-nitzschia* ITS2, the 5.8SF (5'-TGC TTG TCT GAG TGT CTG TGG A-3') and 28SR (5'-TAT GCT TAA ATT CAG CGG GT-3') primer set was used (Lim et al. 2018).

Each reaction was performed with an initial denaturing step at 95°C for 1 minute followed by 30 cycles of 95°C for 15 seconds, 56°C for 15 seconds, and 72°C for 30 seconds. 2.5 µL of each PCR reaction was run on a 1.8% agarose gel to confirm amplification, then PCR products were purified with Beckman Coulter AMPure XP beads following the manufacturer's instructions. PCR quantification was performed in duplicate using the Invitrogen Quant-iT PicoGreen dsDNA Assay kit. Samples were then combined in equal proportions into multiple pools followed by another 0.8x AMPure XP bead purification on the final pool. DNA quality of each pool was evaluated on an Agilent 2200 TapeStation, and quantification was performed with the Qubit HS dsDNA kit. Each 16S or 18S pool was sequenced on an Illumina MiSeq (2 x 300 bp for 16S and V4 or 2 x 150 bp for V9) except for the one pool for the 2014-2016 euphotic zone V9 samples, which was run on an Illumina NextSeq (Mid Output, 2 x 150 bp). The Pseudo-nitzschia ITS2 pool was sequenced on a single lane of an Illumina NovaSeq 6000 with a SP flow cell (2 x 250 bp).

2.5 Analysis of amplicon sequence data

Amplicons were first analyzed with QIIME2 v2019.10 for the 16S and 18S data or QIIME2 v2021.2 for the Pseudo-nitzschia ITS2 data (Bolyen et al. 2019). Briefly, paired-end reads were trimmed to remove adapter and primer sequences with cutadapt (Martin 2011). Trimmed reads were then denoised with DADA2 to produce amplicon sequence variants (ASVs). Each MiSeq run was denoised with DADA2 separately to account for different error profiles in each run then merged. Taxonomic annotation of ASVs was performed with the q2-feature-classifier naïve bayes classifier using the SILVA database (Release 138) for 16S ASVs and the PR2 database (v4.13.0) for 18S ASVs (Bokulich et al. 2018; Guillou et al. 2012;

Pedregosa et al. 2011; Pruesse et al. 2007). For ITS2 ASVs, annotation was performed using BLAST (classify-consensus-blast) with Pseudo-nitzschia ITS2 sequences database as references (Brunson et al. 2024; Lim et al. 2018). When examining 16S relative abundances, all eukaryotic, plastid, and mitochondrial ASVs were removed, and when examining 18S relative abundances, dinoflagellates exclude Syndiniales and prymnesiophytes refers to all Prymnesiophyceae unless otherwise noted.

To estimate absolute abundances of 16S and 18S ASVs, recovery of the aforementioned internal standards was used (Lin et al. 2019). For each ASV within each sample, the number of reads was divided by the ratio of *T. thermophilus* or *S. pombe* reads to the number of rRNA copies added. The total number of copies was then normalized to the volume filtered for each sample to estimate copies L⁻¹.

2.6 Metatranscriptome sequencing, assembly, and analysis

To examine eukaryotic sequences, poly-A selected RNA-Seq libraries were created. Poly-A selected cDNA from total RNA was generated with the SMART-Seq® v4 Ultra® Low Input RNA Kit for Sequencing (Takara Bio USA, Inc.), which was then sheared with a Covaris® E210 focused-ultrasonicator targeting 300 bp fragments. The final sequencing library was then constructed with the NEB NEBNext® Ultra™ II DNA Library Kit and sequenced on three lanes of a NovaSeq 6000 with a S4 flow cell (2 x 150 bp).

For prokaryotic sequences, sequencing libraries were constructed following ribosomal RNA (rRNA) depletion. For samples from 2014-2019, rRNA was depleted with a 2:1:1 mixture of Ribo-Zero Plant, Bacteria, and Human/Mouse/Rat (Illumina) following the manufacturer's low input protocol. For samples from 2020, rRNA depletion was performed with siTOOLS

riboPOOLs with a 6:1:1 mixture of the pan-prokaryote, pan-plant, and pan-mammal riboPOOLs following the manufacturer's instructions for low inputs. rRNA depletion was confirmed with all samples on an Agilent TapeStation 2200 with High Sensitivity RNA ScreenTape. cDNA was then synthesized using the Ovation RNA-Seq System V2 kit (NuGEN) using both poly-A and random hexamer primers followed by fragmentation with a Covaris® E210 focused-ultrasonicator targeting 300 bp fragments. cDNA was then purified with Agencourt RNAClean XP beads, and library preparation was performed with the Ovation Ultralow System V2 (NuGEN). After end repair, ligation, and amplification, the libraries were assessed on an Agilent TapeStation 2200 with High Sensitivity DNA ScreenTape. A subset of samples from 2014-2016 were sequenced separately in four pools on an Illumina HiSeq 4000 (2 x 150 bp). The remaining libraries were sequenced across five pools on an Illumina NovaSeq 6000 with a S4 flow cell (2 x 150 bp).

For both types of libraries, the resulting raw reads were trimmed for quality and to remove Illumina adaptors. Ribosomal RNA sequences were also removed with Ribopicker v0.4.3 (Schmieder et al. 2011). Trimmed and filtered reads were then used for assembly into contigs, and abundances were quantified by mapping these reads to the assembly. Both assembly and read mapping were performed with CLC Bio Genomics Server v21.0.3. Gene prediction was performed with FragGeneScan v1.16, and rRNA removal was performed again with Ribopicker (Rho et al. 2010). Predicted proteins were further filtered to remove those less than 10 amino acids long or with greater than or equal to 20% stop codons. Gene clusters were generated from the predicted proteins with MCL with the inflation option (-I) set to 4 and scheme option (-scheme) set to 6 (Enright et al. 2002).

With the final assemblies, taxonomic annotation of each protein coding gene was assigned via DIAMOND BLASTP alignments using PhyloDB v1.076 as a reference database (Bertrand et al. 2015). Final taxonomic assignments were based on highest Lineage Probability Index values (Podell and Gaasterland 2007). For functional annotation, DIAMOND BLASTP alignments were performed with the Kyoto Encyclopedia of Genes and Genomes (KEGG; Release 94.1) (Kanehisa et al. 2017). KEGG Ortholog (KO) assignment was performed with KofamKOALA, which uses hmmsearch against KOfam, an HMM database of KOs (Aramaki et al. 2019). E-value cutoffs of $1e-5$ were used throughout.

To normalize read counts, raw reads were first converted to transcripts per million (TPM)(Li et al. 2009). Absolute quantities (transcripts L^{-1}), i.e. quantitative metatranscriptomics, were then calculated based on the average recovery of the aforementioned Invitrogen™ ArrayControl™ RNA Spikes and the volume of sample filtered as described in Cohen et al. (2022). For performing correlations with individual genes, transcripts L^{-1} within each taxonomic group and sample were summed by KO annotation, and if KO annotation was absent, the KEGG gene annotation was used. The domoic acid biosynthesis gene, *dabA*, was annotated using the sequences reported in Brunson et al. (2018).

2.7 Flow Cytometry

Two mL samples of seawater were collected in cryovials and fixed with 100 μ L of 0.2 μ m filtered 10% paraformaldehyde for a 0.5% final concentration. Fixation was allowed to occur for 10 min before freezing the sample in liquid nitrogen. Once onshore, samples were stored at -80°C or on dry ice until analysis at the SOEST Flow Cytometry Facility, University of Hawaii at Manoa.

Prior to analysis, thawed samples were stained with Hoechst 33342 (1 $\mu\text{g mL}^{-1}$, v/v, final concentration) at room temperature in the dark for 1 hour (Monger and Landry 1993). Aliquots (100 μl) were analyzed using a Beckman-Coulter EPICS Altra flow cytometer with a Harvard Apparatus syringe pump for volumetric sample delivery. Simultaneous (co-linear) excitation of the plankton was provided by two argon ion lasers, tuned to 488 nm (1 W) and the UV range (200 mW). The optical filter configuration distinguished populations based on chlorophyll *a* (red fluorescence, 680 nm), phycoerythrin (orange fluorescence, 575 nm), DNA (blue fluorescence, 450 nm), and forward and 90° side-scatter signatures. Blue-fluorescence and red-fluorescence signals are used to distinguish DNA-containing heterotrophic (non-pigmented) from phototrophic (chlorophyll-containing) cells. Standardized fluorescence and scatter parameters were determined relative to 0.5- and 1.0- μm yellow-green calibration beads and 0.5- μm UV calibration beads run in each sample. Bead-normalized red-fluorescence values are also used as a measure of cellular chlorophyll *a* (Landry et al. 2003). Raw data (listmode files) were processed using the software FlowJo (Treestar Inc., www.flowjo.com). Cell abundance estimates were then calculated accounting for the volume dilutions from the added preservative and stain solutions.

2.7 Statistics

All correlations and models were generated with R v4.3.2. Generalized additive models (GAMs) were fit with the *mgcv* package v 1.9-1 (Wood 2017). When examining multiple gene correlations, P-values were adjusted using the Benjamini & Hochberg procedure (Benjamini and Hochberg 1995).

3 Results and Discussion

3.1 Quantitative relationships among pigments, DNA, and RNA

As the sampling locations span nearshore stations influenced by coastal upwelling to offshore oligotrophic conditions (James et al. 2022; Venrick 2002), the paired phytoplankton pigment and DNA or RNA samples encompass a wide range of environmental conditions (Figs. 3.1A-3.1D and Supplemental Fig. 3.1). Seasonal and interannual variability were also captured with notably lower temperatures and higher nutrients in the Spring season relating to seasonal upwelling as well as an unprecedented marine heatwave from mid-2014 to 2016 associated with El Niño in 2015 (Di Lorenzo and Mantua 2016; Jacox et al. 2018). Concurrent with the wide-ranging environmental conditions, total chlorophyll *a* and accessory pigment concentrations often spanned several orders of magnitude with all accessory pigments detected in concentrations as low as 0.001 to 0.005 $\mu\text{g L}^{-1}$ (Fig. 3.1E). Fucoxanthin (Fuco) exhibited the greatest range with concentrations reaching 6.81 $\mu\text{g L}^{-1}$, and 19'-hexanoyloxyfucoxanthin (HexFuco) exceeded 1 $\mu\text{g L}^{-1}$ in several samples. All other pigment concentrations examined here were always less than 1 $\mu\text{g L}^{-1}$.

DNA metabarcoding of both the V4 and V9 regions of the 18S rRNA gene for eukaryotes as well as the V4-V5 regions of the 16S rRNA gene for prokaryotes similarly showed wide ranges in relative abundances (Figs. 3.1F and 3.1G). Among eukaryotes, diatoms exhibited the greatest range in accordance with their known dominance during blooms in the region (Goericke 2011; Venrick 2012). Chlorophytes also exhibited a wide range of relative abundances, and their average abundance exceeded that of diatoms. Dinoflagellates were also found in high abundances, and while they indeed bloom in the region (Anderson et al. 2008; Kahru et al. 2021), their high abundances via 18S metabarcoding may be due to bias from high 18S gene

copy numbers (Martin et al. 2022). Other eukaryotic phytoplankton groups were typically less than 10% of the total eukaryotic community. In particular, cryptophytes were less than 1% of the community on average but were found to be as high as 5%. Within eukaryotic mRNA expressed as relative abundances [transcripts per million (TPM)], intragroup variability was less pronounced, and most groups displayed higher relative abundances compared to their 18S DNA abundances (Figs. 3.1F). The sole exception was with dinoflagellates which had lower relative abundances with RNA compared to DNA, further suggesting bias in their DNA relative abundances due to high 18S DNA copy numbers.

Within prokaryotes, all cyanobacteria combined and the cyanobacterial genus *Prochlorococcus* displayed similar distributions, indicating that *Prochlorococcus* was often the dominant member of the cyanobacterial community (Figs. 3.1G and 3.1I). Cell abundances measured via flow cytometry also show that *Prochlorococcus* is dominant in these samples, with *Prochlorococcus* displaying higher abundances than *Synechococcus* in 76% of samples (Supplemental Fig. 3.2). Cyanobacteria as a whole and *Prochlorococcus* also displayed lower relative abundances from transcripts compared to 16S rRNA genes (Fig. 3.1G).

When converted to absolute abundances based on the recovery of internal standards and normalizing to the volume filtered (gene copies L⁻¹ or transcripts L⁻¹), differences and distributions among groups were similar to those observed with relative abundances (Figs. 3.1F-1I). However, such quantitative approaches have not been widely adopted, and most metabarcoding data are solely expressed as relative abundances. As previously described, other studies have normalized phytoplankton pigments to total chlorophyll *a* concentrations in an attempt to make DNA-based relative abundances and phytoplankton pigments comparable (Catlett et al. 2023; Kramer et al. 2024a). When using pigments directly, diagnostic pigments

used for specific phytoplankton groups include monovinyl chlorophyll *b* (MVChlb) for chlorophytes, alloxanthin (Allo) for cryptophytes, fucoxanthin (Fuco) for diatoms, peridinin (Perid) for dinoflagellates, 19'-butanoyloxyfucoxanthin (ButFuco) for pelagophytes, 19'-hexanoyloxyfucoxanthin (HexFuco) for prymnesiophytes, zeaxanthin (Zea) for cyanobacteria, and divinyl chlorophyll *a* (DVChla) for *Prochlorococcus* (Catlett et al. 2023; Kramer and Siegel 2019).

Within eukaryotic taxa, comparisons between normalized pigment concentrations and DNA-based relative abundances showed moderate to strong correlations for several groups (Fig. 3.2A). In particular, the relative abundances of diatoms and Fuco displayed the strongest agreement ($r = 0.79-0.82$) followed by pelagophytes and ButFuco ($r = 0.71-0.77$). Cryptophytes with Allo and chlorophytes with MVChlb were also moderately well correlated ($r = 0.51-0.65$) whereas dinoflagellates with Perid and prymnesiophytes with HexFuco were weakly correlated ($r = 0.32-0.42$). Previous studies using samples from the California Current (Plumes and Blooms), North Atlantic Ocean (NAAMES), and Northeast Pacific Ocean (EXPORTS) have shown similar results where relationships are notable for chlorophytes, cryptophytes and pelagophytes, weak for prymnesiophytes and dinoflagellates, and inconsistent among the studies for diatoms (Catlett et al. 2023; Kramer et al. 2024a). In another previous study in the Western Antarctic Peninsula where a chemotaxonomic approach that partitions total chlorophyll *a* (CHEMTAX) was used rather than diagnostic pigments directly, cryptophytes were moderately well correlated between the measurements whereas diatoms were not (Lin et al. 2019; Mackey et al. 1996). Within these studies, abundances for some groups were also too low to be included.

Although these results suggest that there are moderately strong relationships between pigments and DNA-based relative abundances for several phytoplankton groups, these analyses

are hindered by the compositional nature of relative abundance data as previously described. The estimation of absolute abundances via internal standards removes compositionality and enables the comparison of DNA- or RNA-based concentrations (copies L⁻¹ or transcripts L⁻¹) to phytoplankton pigment concentrations (µg L⁻¹). By doing so here, all correlations were stronger than those with relative abundances for eukaryotic phytoplankton taxa (Fig. 3.2B). In particular, the relationships between dinoflagellates and Perid was dramatically stronger with both marker genes ($r = 0.77-0.79$). The relationship for prymnesiophytes was also stronger but not to the same degree ($r = 0.46-0.51$). Furthermore, all groups were most strongly correlated with their diagnostic pigments except prymnesiophytes, which had similarly strong correlations with peridinin (Supplemental Fig. 3.3). By using taxon-specific biomass as chlorophyll *a* concentrations with a chemotaxonomic approach (phytclass) (Hayward et al. 2023), correlations were often similar (Supplemental Fig. 3.4). Moreover, strong correlations were also observed among absolute transcript abundances and pigment concentrations ($r = 0.66-0.88$) except for prymnesiophytes which had a similar correlation when using their DNA-based abundances ($r = 0.43$) (Fig. 3.2C).

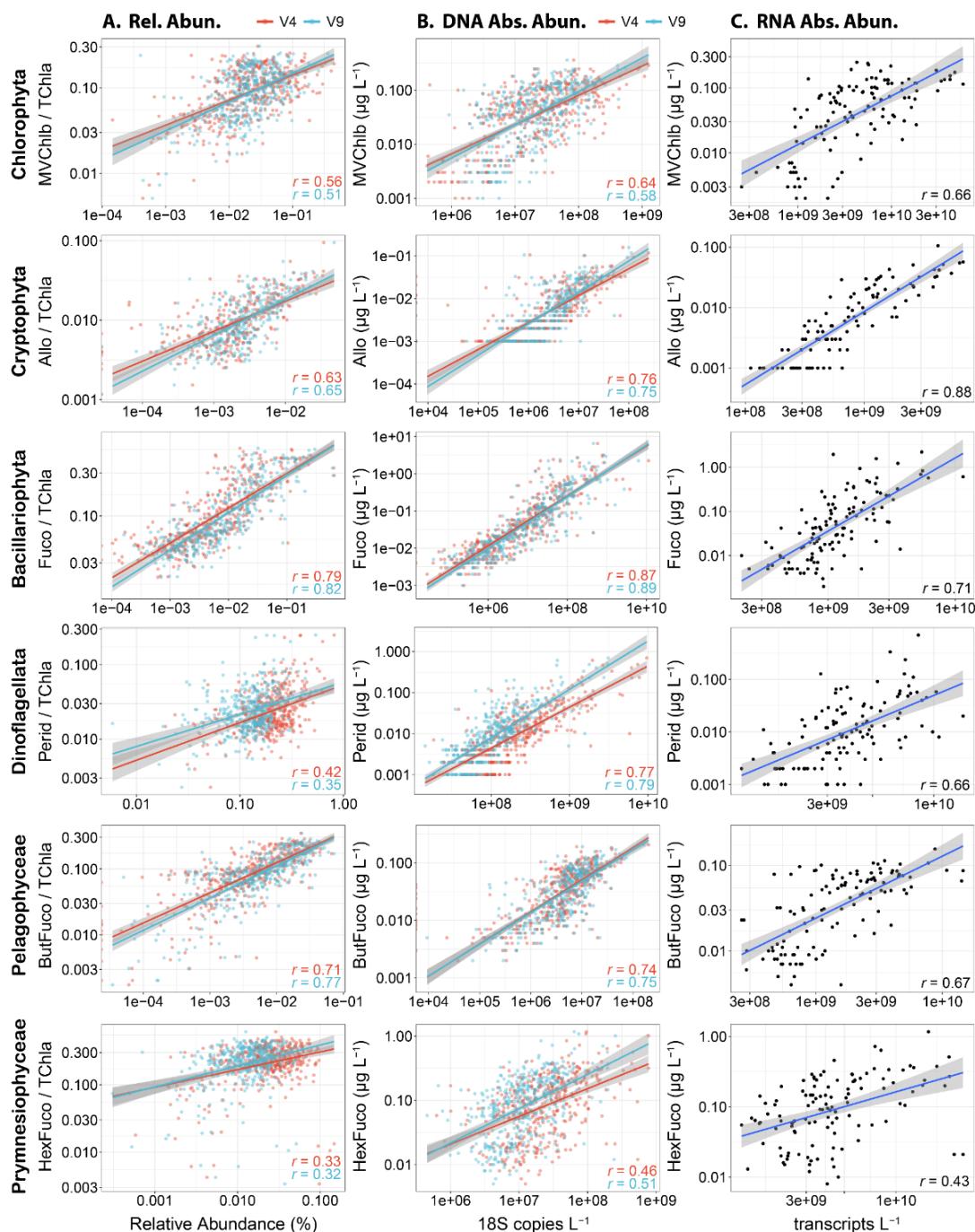


Figure 3.2: Correlations among abundance metrics for each major eukaryotic phytoplankton group. Pearson correlations are all significant ($P < 0.05$) and coefficients are displayed in each panel. Lines with 95% confidence intervals show linear models with significant relationships between variables ($P < 0.05$). (A) Diagnostic pigments normalized to total chlorophyll *a* (TChla) against relative abundances of either the 18S-V4 (red) and 18S-V9 (blue) rRNA gene abundance. (B) Diagnostic pigment concentrations against absolute abundances of the 18S-V4 (red) and 18S-V9 (blue) rRNA genes (copies L⁻¹). (C) Diagnostic pigment concentrations against absolute abundances of mRNA (transcripts L⁻¹).

These results indicate that abundances of diagnostic phytoplankton pigments and phytoplankton DNA or RNA often agree. Furthermore, they suggest that the aforementioned issues with compositionality or normalization of pigments to total chlorophyll *a* may lead to misleading conclusions about these relationships, particularly for dinoflagellates. Besides compositionality, DNA- and RNA-based abundances have other biases that can also lead to discrepancies. Copies of 18S DNA can vary orders of magnitude among closely related phytoplankton species (Zhu et al. 2005), and DNA sequencing may also capture deceased cells that have lost some or all of their pigments. RNA abundances may be variable within a cell, instead reflecting overall activity or differential transcription rates in response to environmental stimuli (Moran et al. 2012). Both metrics are also influenced by the comprehensiveness of reference databases where unknown phytoplankton sequences may be unassigned (Krinos et al. 2023).

Differences in pigment abundances may also be a source of variability. For example, reduced light availability may lead to cellular increases in accessory pigments (Henriksen et al. 2002). Using mixed layer depth as a proxy for light history, decreased light availability has been shown to have a small but significant effect on discrepancies between phytoplankton pigments and DNA relative abundances (Catlett et al. 2023). To examine the effects of light history, we performed linear regressions on the residuals from linear models for absolute abundances with depth as a predictor variable and coarse proxy for reduced light availability (Supplemental Fig. 3.5). Besides with cryptophytes, depth was found to positively and significantly predict residuals, aligning with an increase in pigments at depth which may be attributed to reduced light. High variability in near-surface samples was also observed and may reflect large differences in light conditions from sampling throughout the diel cycle.

Furthermore, as some pigments are shared among certain groups, the presence of one group may influence pigment concentrations that are diagnostic of a different group.

Fucoxanthin, for example, is not only found in diatoms but also other photosynthetic eukaryotes with red algal-derived plastids (Jeffrey et al. 2011; Kramer and Siegel 2019). Previous studies suggest that shared pigments are an important source of disagreement between the two measurements, but the strong correlations for many groups, particularly with diatoms, suggests that it may be a less prevalent issue in certain cases (Catlett et al. 2023; Kramer et al. 2024a).

Intragroup variability in pigment composition may also be a contributing factor with different species having different pigment concentrations (Neeley et al. 2022; Zapata et al. 2004). In particular, dinoflagellates may be influenced by taxa that no longer have peridinin. The previously shown dinoflagellate correlations were performed without Syndiniales, an early branching clade that are likely parasitic and have lost their plastid altogether (Decelle et al. 2022; Guillou et al. 2008). Overall, Syndiniales accounted for 30 to 34% of dinoflagellate 18S copies in our data, and their removal considerably increased the strength of dinoflagellate-peridinin correlations (Supplemental Fig. 3.6 and Supplemental Table 3.1). Other dinoflagellate taxa appear to have instead replaced their plastids with others via kleptoplasty, integration of a new plastid, or obtaining endosymbionts (Novák Vanclová and Dorrell 2024). For example, the Kareniaceae family (e.g., the genera *Karenia*, *Karlodinium*, and *Takayama*) and Kryptoperidinaceae family (e.g., the genus *Durinskia*) contain fucoxanthin from an ancestral haptophyte or diatom endosymbiont respectively, members of the genus *Dinophysis* have a complex kleptoplastidic relationship with a cryptophyte, and the genus *Lepidodinium* contains plastids originating from a chlorophyte (Kamikawa et al. 2015; Novák Vanclová and Dorrell 2024; Zapata et al. 2012). However, neither Kryptoperidinaceae nor *Lepidodinium* were detected

in these data with either 18S marker, and filtering of other groups did not alter the strength of the correlation between Dinoflagellates and peridinin indicating that besides Syndiniales, these groups do not influence the dinoflagellate-peridinin relationship in this region (Table S1). The prymnesiophyte *Phaeocystis globosa* has also been suggested to contain ButFuco rather than HexFuco (Antajan et al. 2004; Wang et al. 2022), and while filtering ASVs annotated as this species marginally improved the strength of the correlation for prymnesiophytes with the V4 data (from $r = 0.46$ to $r = 0.48$), it was lower in the V9 data (from $r = 0.51$ to $r = 0.45$).

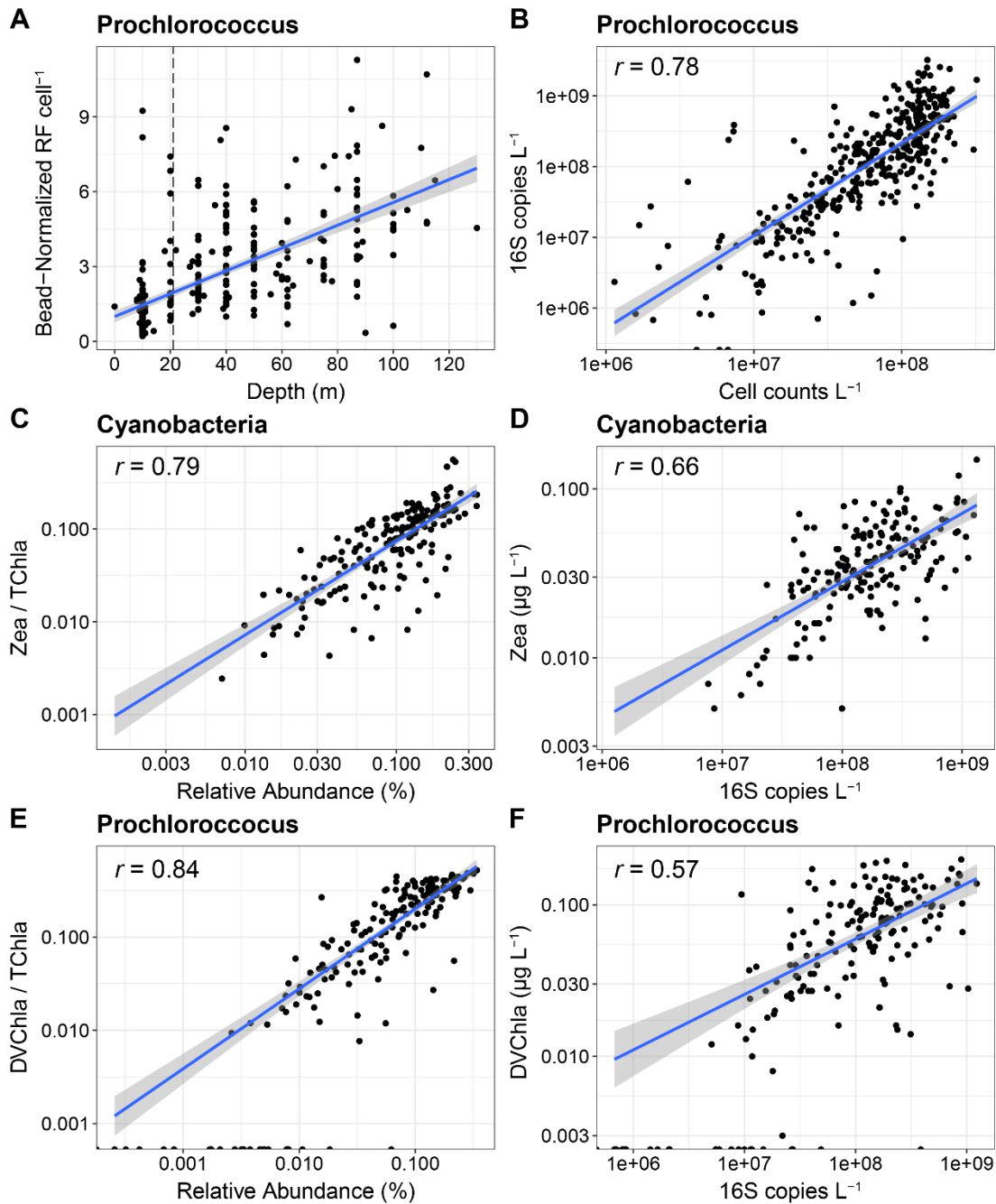


Figure 3.3: Correlations among abundance metrics for cyanobacteria. (A) Average bead-normalized red fluorescence per cell for *Prochlorococcus* versus depth (m). The fluorescence units are arbitrary. (B) *Prochlorococcus* DNA-based absolute abundances (copies L⁻¹) against cellular abundances. (C, D) Relative and absolute abundance comparisons between all cyanobacteria and zeaxanthin (Zea). (E, F) Relative and absolute abundance comparisons between *Prochlorococcus* and divinyl chlorophyll *a* (DVChla). Pearson correlations are significant ($P < 0.05$) and coefficients are displayed in panels B-F. Lines with 95% confidence intervals shows linear models with significant relationships between variables ($P < 0.05$).

Within cyanobacteria, *Prochlorococcus* chlorophyll content significantly increased with depth, suggesting an increase in pigments as light availability is reduced (Fig. 3.3A). Furthermore, *Prochlorococcus* absolute abundances strongly agreed with cell abundances from flow cytometry throughout the water column as previously observed in the greater region (Fig. 3.3B) (Jones-Kellett et al. 2024). In combination, these results suggest that DNA-based abundances and pigments are increasingly uncoupled as light availability is reduced. To minimize effects from light, only samples from the upper 20 m of the water column were further examined. When using relative abundances, all cyanobacteria with *Zea* and *Prochlorococcus* with DVChla were strongly correlated (Fig. 3.3C-F). However, DVChla, was not detected in 12.5% of samples which generally corresponded to lower DNA abundances (Fig. 3.3F). Unlike the eukaryotic data, the strengths of these correlations were lower when using absolute abundances, albeit still moderately well correlated. Although variability in pigments was reduced by removing samples deeper than 20 m, other factors such as cloud cover may further contribute the remaining observed pigment variability and a weaker correlation with absolute abundances. This variation in pigments may be normalized when dividing by total chlorophyll *a* leading to higher correlations with relative abundances. Some of the discrepancy with absolute abundances may also be caused by *Prochlorococcus* cell sizes which are smaller than the pore size of the GF/F filters used for pigments here (nominally 0.7 μm), leading to some *Prochlorococcus* cells being missed while being captured by flow cytometry or DNA that used a smaller pore size (0.2 μm) (Partensky et al. 1999; Ting et al. 2007). As in eukaryotic phytoplankton, the absolute abundances of transcripts were also moderately well correlated for cyanobacteria with *Zea* and *Prochlorococcus* with DVChla (Supplemental Fig. 3.7).

3.2 Applications from integration phytoplankton pigments with molecular data

The strong agreement among eukaryotic phytoplankton pigments and DNA or RNA for several groups indicates that these metrics are comparable proxies for phytoplankton abundances and community composition. As phytoplankton pigments directly impact remote sensing reflectance spectra, HPLC pigments are useful for validating remote-sensing algorithms; however, these results suggest that the absolute abundances of DNA or RNA may also be useful for model development. Furthermore, these results support that potential models for phytoplankton pigment concentrations via remote sensing may be able to provide comparable global estimates of different phytoplankton groups (Kramer et al. 2022). In the following sections, we further examine relationships gleaned by integrating phytoplankton pigments and DNA- or RNA-based metrics to demonstrate potential applications for addressing ecological questions, monitoring harmful algal blooms, or inferring phytoplankton group-specific activity.

3.2.1 Increased taxonomic resolution with biomass estimation in ecological assessments

Phytoplankton pigments and DNA-based metabarcoding have separate strengths that can be leveraged when integrated: phytoplankton contributions to total chlorophyll *a* can be estimated with chemotaxonomic approaches providing a standardized estimate of biomass for each group, and DNA offers marker gene-level resolution into the composition of each sample. This increased resolution is possible even when the DNA-based data are expressed solely as relative abundances.

To illustrate this combined approach, diversity expressed as the Shannon index (H') was compared to both total chlorophyll *a* and taxon-specific chlorophyll *a* as proxies for phytoplankton biomass (Fig. 3.4). Biomass is often used as a proxy for productivity, thus

enabling investigation of productivity-diversity relationships (PDRs) (Irigoien et al. 2004; Smith 2007).

PDRs may exhibit different trends, but marine phytoplankton are presumed to exhibit a unimodal distribution with maximum diversity at an intermediate level of productivity, including within models of phytoplankton communities in the California Current Ecosystem (Goebel et al. 2013; Irigoien et al. 2004; Li 2002). PDRs may also be positive, negative, or flat, with other studies suggesting that there is no relationship when accounting for potentially inadequate sampling (Cerreño et al. 2013; Smith 2007). Examination of these relationships is particularly important for understanding how environmental change may impact diversity and productivity.

When comparing total chlorophyll *a* concentrations and the diversity of all six eukaryotic phytoplankton groups examined here, the PDR for the region was negative rather than unimodal, with diversity remaining high at low and intermediate biomass levels before declining at high biomass (Fig. 3.4A). During the same cruises, additional samples concurrently measured DNA abundances and net primary productivity (NPP, $\text{mg C m}^{-3} \text{ d}^{-1}$). When comparing diversity and NPP directly, the relationship was similarly flat then negative (Supplemental Fig. 3.8A). The Shannon index considers both richness and evenness, thereby downweighing the influence of rare taxa in comparison to richness, defined here as the number of ASVs (Ibarbalz et al. 2019; Ma 2018). Despite substantially more variability, richness also displayed a negative trend including when adding cyanobacterial richness (Supplemental Fig. 3.8B and 3.9).

By partitioning total chlorophyll *a* concentrations into separate eukaryotic phytoplankton groups and leveraging the high taxonomic resolution of the DNA-based data, PDRs were examined within each group (Fig. 3.4). Dinoflagellates, diatoms, and cryptophytes displayed significant relationships whereas other groups did not for both amplicons. Furthermore, diatoms

and dinoflagellates displayed opposing trends, where diatoms displayed the expected unimodal relationship while dinoflagellates displayed a negative relationship. When considering richness, the trend for diatoms was largely positive, indicating a decline in evenness at the highest levels of biomass leading to lower diversity, and for cryptophytes, the relationship was slightly positive with both metrics (Fig. 3.4 and Supplemental Fig. 3.10). Dinoflagellates had the highest richness with as many as 353 ASVs, while diatoms had a maximum of 93 ASVs (Supplemental Fig. 3.10). Cryptophytes and pelagophytes had low richness with maxima of 15 and 9 ASVs, respectively (Supplemental Fig. 3.10). For pelagophytes, the relatively low richness and lack of trend may be due to dominance of a single species, *Pelagomonas calceolata*, that is highly prevalent in SCMLs (Dupont et al. 2015; Guérin et al. 2022). Meanwhile, chlorophytes and prymnesiophytes did not display significant relationships between biomass and richness. Considering the relatively high richness and sole negative trend, dinoflagellates appear to be the primary drivers of the community-wide trend, besides at high biomass where diatoms have a high influence. This dominance by dinoflagellates is further evidenced by a shift to a more unimodal distribution by examining the community in the absence of dinoflagellates (Supplemental Fig. 3.11).

The observations of unimodal PDRs have led to hypotheses for the mechanisms that underlying them. Indeed, our observed negative trend aligns with the negative side of unimodal PDRs where diversity decreases with increased productivity. This decline is predicted to be associated with high productivity nearshore upwelling conditions where there is strong competition for light and opportunist large cells such as diatoms escape grazing (Goebel et al. 2013; Irigoien et al. 2004; Vallina et al. 2014). In agreement, the observed low diversity aligns with the nearshore environment, shallow nitracline depths, and increase in diatom biomass

relative to dinoflagellates (Figs. 3.4 and Supplemental Fig. 3.12). However, increased diatom richness with biomass indicates that the negative trend is not entirely driven by the dominance of a few opportunist diatom taxa that escape predation (Supplemental Fig. 3.10C). Rather, many diatoms appear to flourish when competition for nutrients is minimized under upwelling conditions, although evenness likely declines at the highest levels.

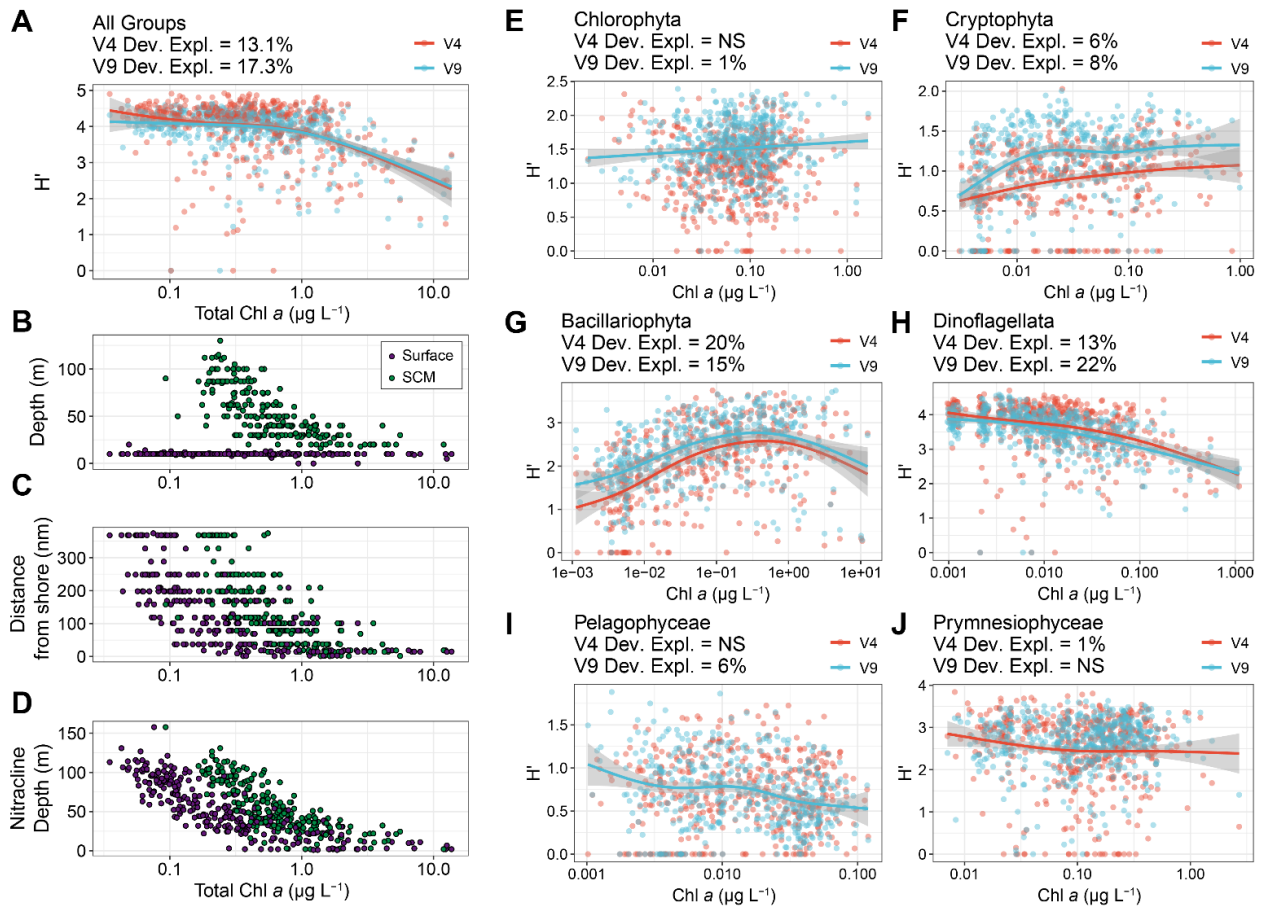


Figure 3.4: Productivity-diversity relationships in the region with biomass (chlorophyll *a* concentrations) as a proxy for productivity and diversity expressed as the Shannon Index (H') for both the 18S-V4 (red) and 18S-V9 (blue) data. (A) The diversity of all eukaryotic phytoplankton groups and total chlorophyll *a* concentrations. (B-D) Environmental variables against total chlorophyll *a* concentrations colored by near-surface (purple) or subsurface chlorophyll max (SCM, green) samples. (E-J) Diversity of individual phytoplankton groups against their taxon-specific biomass estimated with phytoclass. Lines represent GAMs and corresponding 95% confidence intervals where significant ($P < 0.05$). The deviance explained by each GAM or “NS” for not significant ($P > 0.05$) is shown above each panel for each amplicon.

High diversity at intermediate productivity has been suggested to be associated with offshore oligotrophic conditions (Goebel et al. 2013). Although offshore oligotrophic samples aligned with the lowest biomass levels, they were in part responsible for intermediate NPP along with SCMLs at intermediate depths, in agreement with model predictions (Fig. 3.4 and Supplemental Fig. 3.8). The low productivity and low diversity end of the unimodal distribution is expected to be caused by light limitation (Goebel et al. 2013), and while chlorophyll *a* concentrations suggest that the deepest SCML samples align with intermediate biomass, they are influenced by reduced light causing elevated chlorophyll content (Cullen 2015). As predicted, the deepest SCML samples displayed the lowest NPP rates; however, contrary to predictions, diversity and richness remained high in these samples resulting in an absence of the positive side of a unimodal distribution (Supplemental Fig. 3.8).

The low productivity and diversity end of unimodal distributions have also been attributed to selective grazing with the dominance of a few slow-growing nutrient specialists (Vallina et al. 2014). As diversity and richness instead remained high, many phytoplankton taxa, particularly dinoflagellates, appear to coexist within low productivity regimes. Within dinoflagellates, this coexistence may be supported by mixotrophy or diel vertical migrations where nutrient availability at depth is exploited at night and photosynthesis in the near-surface occurs during the day (Stoecker et al. 2017; Zheng et al. 2023). Within prymnesiophytes, there may also be some mixotrophic taxa (Koppelle et al. 2022). Small cells such as chlorophytes, pelagophytes, and prymnesiophytes are also at an advantage under oligotrophic conditions due to more effective resource acquisition and use (Raven 1998), although their diversity is maintained across their biomass ranges (Fig. 3.4). Even though diatom diversity is lower under oligotrophic conditions, certain diatoms such as those that form symbiotic relationships with diazotrophs may

be favored under these conditions and contribute to the increased diversity (Kemp and Villareal 2018).

Overall, these results highlight the contrasting strategies of different phytoplankton groups and align with the classical view of diatoms and dinoflagellates on opposite ends of the r-selected vs K-selected continuum (Margalef 1978). However, these results may differ from other regions that exhibit greater stability or experience frequent blooms of groups besides diatoms. Continued warming in the region is anticipated to lead to increased stratification resulting in conditions analogous to those in the offshore oligotrophic region with deeper nitraclines (Ducklow et al. 2022; Lund 2024). These projections imply that increased stratification leads to low productivity but high phytoplankton community diversity, although diatom diversity will be lower. In such a scenario, the high level of diversity will instead be driven by dinoflagellates and supplemented by a variety of picoeukaryotes, contrasting with predictions of lower diversity in ecosystem models of open ocean regions (Henson et al. 2021).

3.2.3 Monitoring and forecasting harmful algal blooms

The strong relationships between pigments and abundances of certain groups may also be useful for monitoring or forecasting harmful algal blooms (HABs). For example, certain species in the diatom genus, *Pseudo-nitzschia*, produce the neurotoxin domoic acid (DA) resulting in HABs worldwide (Bates et al. 2018). They are also commonly responsible for HABs in our study region and California Current at large (Lewitus et al. 2012). With absolute abundances accounting for 10% to 11% of diatom 18S copies, *Pseudo-nitzschia* was among the most dominant diatom genera, only exceeded by *Thalassiosira* and *Chaetoceros* (Fig. 3.5A). *Pseudo-nitzschia* was also detected in 74% to 79% of samples when diatoms were present with an

equivalent number when fucoxanthin was detected. Overall, a strong positive relationship between *Pseudo-nitzschia* and total diatom abundances was observed, and fucoxanthin concentrations explained 10% more of the variance in *Pseudo-nitzschia* abundances than chlorophyll *a* concentrations (Fig. 3.5). Expression of *dabA*, the first gene in the domoic acid biosynthetic pathway (Brunson et al. 2018), was also detected in 13 metatranscriptomics samples, and 7 of the 11 samples with greater than 0.5 µg L⁻¹ concentrations of fucoxanthin had detectable *dabA* expression (Fig. 3.5E). These samples comprised diverse *Pseudo-nitzschia* species with an overall dominance of *P. delicatissima*; however, the highest *dabA* expression occurred when *P. australis* relative abundances were elevated (Fig. 3.5F).

Not all species of *Pseudo-nitzschia* have been shown to produce domoic acid, and DA production is influenced by environmental conditions (Bates et al. 2018). Similarly, expression of *dabA* does not always confer detectable particulate DA (Brunson et al. 2024). However, current models in the region that predict *Pseudo-nitzschia* HABs and domoic acid production use a variety of data including remotely-sensed chlorophyll *a* and two reflectance wavebands, suggesting that fucoxanthin detection in conjunction with other measurements and cellular modeling may offer better predictions for *Pseudo-nitzschia* and domoic acid (Anderson et al. 2016; Moreno et al. 2022).

Certain dinoflagellates may also cause HABs globally and in the region (Anderson et al. 2012; Anderson et al. 2021). These HABs are caused by certain species in the genera *Alexandrium*, *Dinophysis*, and *Gonyaulax* as well as the species *Gymnodinium catenatum* and *Lingulodinium polyedra* (Anderson et al. 2021; Ternon et al. 2023; Trainer et al. 2010). These genera were also among the most dominant dinoflagellate genera detected, although 39% of V4

and 55% of V9 18S copies for dinoflagellates were unassigned on a genus level (Supplemental Fig. 13).

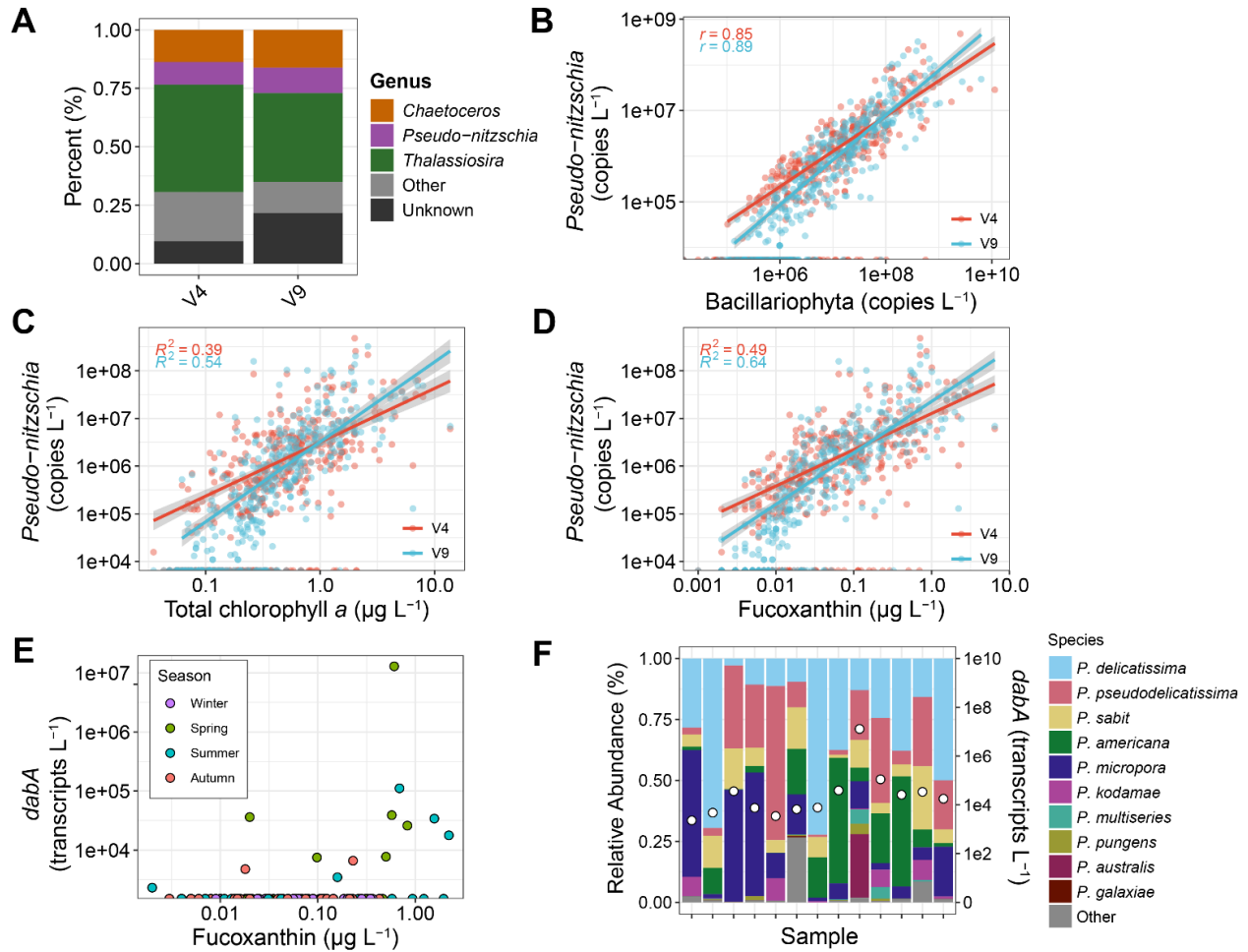


Figure 3.5: *Pseudo-nitzschia* abundances. (A) Relative abundances of the top three diatom genera based on total copies L⁻¹. (B) *Pseudo-nitzschia* absolute abundances against diatom (Bacillariophyta) absolute abundances. (C) *Pseudo-nitzschia* absolute abundances against total chlorophyll *a* concentrations (D) *Pseudo-nitzschia* absolute abundances against fucoxanthin concentrations. (E) Total *dabA* expression against fucoxanthin concentrations. (F) Relative abundances of *Pseudo-nitzschia* species from ITS2 sequencing (left y-axis) and total *dabA* expression (right y-axis). Samples are ordered by fucoxanthin concentrations as shown in Panel E.

Despite dinoflagellate diversity declining with increasing biomass, dinoflagellate richness was the highest among groups, including where the highest dinoflagellate biomass was observed

(Fig. 3.4 and Supplemental Fig. 10). As a result, the samples with high dinoflagellate abundances still comprise many genera. Considering that peridinin concentrations did not exceed $1 \mu\text{g L}^{-1}$, an intense dinoflagellate bloom that may have resulted in even lower diversity was not captured here. Nevertheless, peridinin concentrations and V4-based abundances of the dinoflagellate genera *Alexandrium*, *Gonyaulax*, and *Gymnodinium* were significantly related (Supplemental Fig. 13). Peridinin concentrations also explained 9-16% more of the variance in the abundances of these genera compared to total chlorophyll *a* concentrations. In contrast, no significant relationships were observed for *Lingulodinium* and *Dinophysis*, the latter of which does not contain peridinin as previously noted. As in *Pseudo-nitzschia*, there are non-toxic members of these genera, and their presence does not imply the production of their respective toxins (Anderson et al. 2012). However, the ability to distinguish increased peridinin or fucoxanthin concentrations with remote sensing suggests increases in the genera identified here and may aid HAB monitoring efforts.

3.2.3 Towards increased inference of biogeochemical or metabolic activity

Current ESMs are unable to confidently predict climate-driven changes to NPP, and improving these models to better account for phytoplankton abundances and functions is a critical component to address this uncertainty (Kwiatkowski et al. 2020; Tagliabue 2023). Model parameters include biological rates and biogeochemical fluxes, and there is increasing interest in the ability to connect ‘omics data with rates such that the ‘omics can inform these model parameters (Saito et al. 2024; Strzepek et al. 2022). The absolute quantities of certain proteins have shown promise for inferring rates of nitrite oxidation and carbon fixation (Roberts et al.

2024; Saito et al. 2020), although it is unclear if absolute transcript abundances will be able to serve a similar purpose (McCain et al. 2024).

The strong correlations between pigments and total transcript abundances in most groups examined here suggests that pigment abundances may also relate to the expression of specific genes (Fig. 2C). Within each group, the absolute abundances of genes clustered by KEGG annotations (Sect. 2.6) were correlated with their respective diagnostic pigments (Fig. 3.6). In chlorophytes, there was only one strongly significant correlation with a light-harvesting chlorophyll-binding protein (LHCB) and *Prochlorococcus* which had no significant correlation. In other groups however, 67 to 2,312 genes were strongly correlated ($r > 0.60$, $\text{FDR} < 0.05$). For some taxa, specific genes displayed stronger correlations than both 18S rDNA and all transcripts combined (Fig. 3.2), indicating that they may individually be used as indicators of abundances. These include the pentose phosphate pathway genes 6-phosphogluconate dehydrogenase (*PGD*) and glucose-6-phosphate dehydrogenase (*G6PD*) in dinoflagellates as well as several accessory light harvesting complex (*LHCF*) proteins in prymnesiophytes. With 34% to 65% of genes with no assigned function, the remaining genes fall into diverse sets of metabolic categories (Fig. 3.6).

In cryptophytes and dinoflagellates, several genes related to carbon fixation and photosynthesis were strongly correlated (Fig. 3.6). Cryptophytes and diatoms also showed strong correlations with several nitrogen metabolism genes. In baker's yeast (*Saccharomyces cerevisiae*), the abundances of individual proteins are generally poor predictors of their corresponding reaction rates (McCain et al. 2024), but the combined abundances of functional units of genes, i.e. modules or subsystems, that are responsible for specific pathways may more accurately predict rates. Similarly, increases in the abundances of specific pigments here is indicative of the expression of certain pathways which may be useful in predicting group-

specific reaction rates *in situ*. Further establishment of these relationships or between pigments and protein abundances may potentially support global estimates of these group-specific reaction rates.

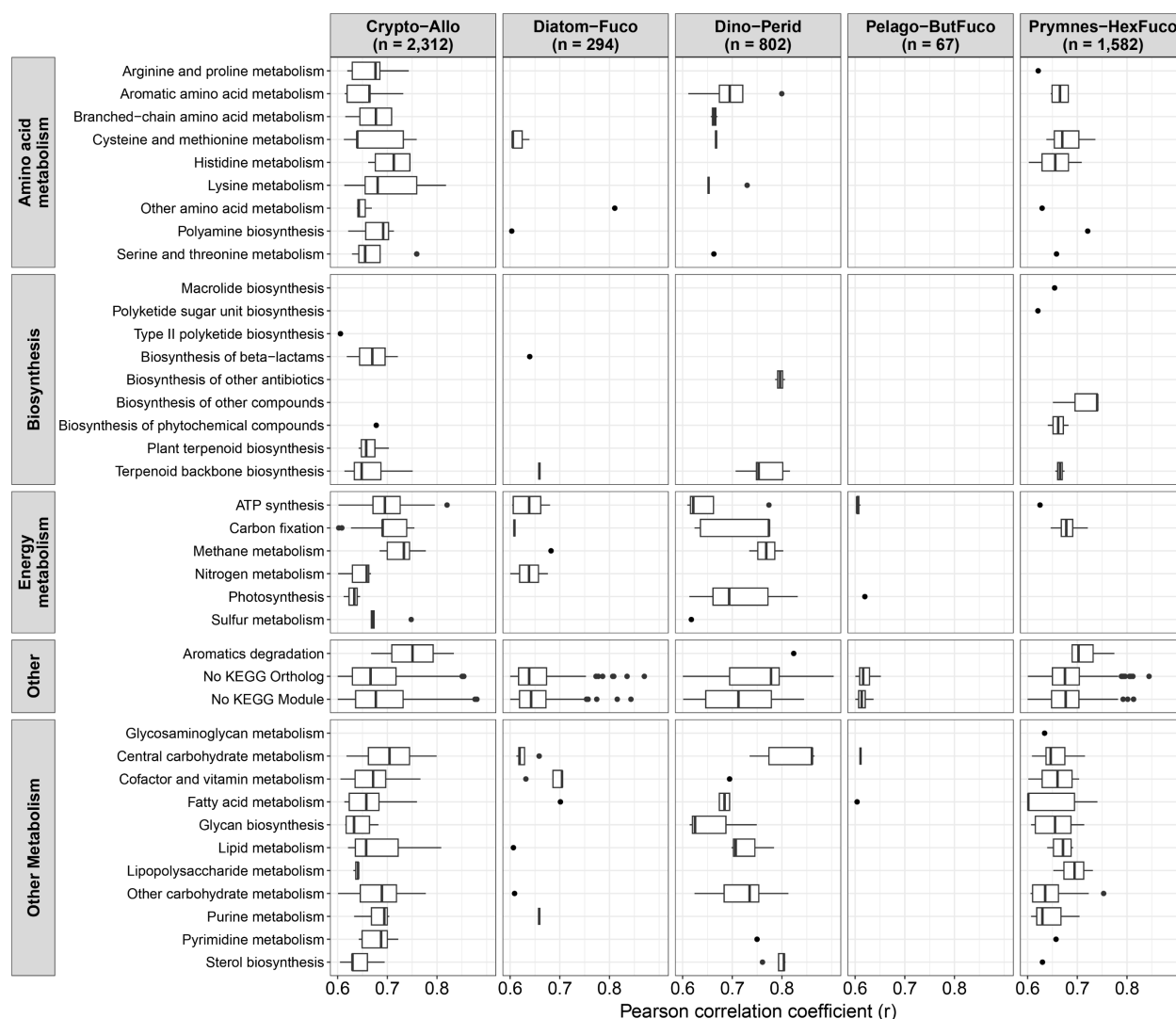


Figure 3.6: Strongly significant correlations between diagnostic pigments and gene expression (transcripts L^{-1}) associated with each respective group ($r > 0.60$, FDR < 0.05) and organized by KEGG module classes. To examine correlations, genes were aggregated by KEGG Ortholog annotations and if not present, the KEGG gene annotations were used.

4 Conclusions

By integrating phytoplankton pigments with quantitative abundances of 18S DNA and total mRNA via metabarcoding and metatranscriptomics respectively, we demonstrate that diagnostic pigments for specific eukaryotic phytoplankton groups correlate with both their DNA- and RNA-based abundances. Although there are inherent biases associated with each of these measurements, their relationships suggest that they are comparable and may be useful for the development of satellite-based remote sensing models of phytoplankton group-specific abundances. Rather than the 18S rRNA gene which suffers from variable copy numbers, other sequencing-based markers may be more useful for drawing these comparisons. For example, the photosystem gene *psbO* is universal among phytoplankton which normally have one or two at most copies per genome (Pierella Karlusich et al. 2023). It is also unclear if 18S sequencing of rRNA rather than the rRNA gene from DNA offers a less biased assessment. However, these results also provide increased confidence that both DNA- and RNA-based abundances are reflective of phytoplankton group-specific biomass. Prymnesiophytes displayed the weakest correlations out of all groups examined here, but the calcium carbonate shells of coccolithophores are highly optically refractive enabling easier detection via satellite-based remote sensing (Balch 2018). When coupled to other satellite, glider, or float-based measurements of photophysiology (Lin et al. 2016; Ryan-Keogh et al. 2023), phytoplankton group-specific abundances, community composition, and physiological assessments may be able to be remotely and collectively assembled.

Although this assessment was only performed in the California Current Ecosystem, these relationships may extend to other regions. Previous use of quantitative metabarcoding in the Western Antarctic Peninsula with relatively high abundances of cryptophytes showed that their

chemotaxonomic abundances were well correlated with their DNA-based abundances (Lin et al. 2019). Diatoms were not strongly correlated, but this result may be an artifact from a small sample size that did not capture a large range of diatom abundances in the region. Considering that the strengths of correlations in all eukaryotic groups improved when translating relative abundances into absolute abundances here, previous studies in the California Current, North Atlantic and Northeast Pacific that used relative abundances were likely affected by issues when correlating compositional data, and we hypothesize that stronger correlations would be observed had quantitative approaches been employed. In particular, dinoflagellates displayed relatively low correlations with relative abundances (Catlett et al. 2023; Kramer et al. 2024a), but the strength of these correlations dramatically improved by using quantitative approaches here (Fig. 3.2A). For cyanobacterial abundances, pigment and transcriptional variability likely contributed to weaker correlations when using absolute abundances. Smaller filter pore sizes such as those obtained by combusting GF/F filters may also need to be considered when drawing these comparisons in future studies (Nayar and Chou 2003).

The existence of relationships among these metrics also opens the door to several applications beyond simply assessing phytoplankton abundances and community composition. For example, PDRs can be examined in detail with not only the whole phytoplankton community but also within individual phytoplankton groups. By doing so, we show that the PDR for the whole community in the region is negative rather than the expected unimodal distribution (Fig. 3.4A). By partitioning biomass into separate groups and leveraging the high resolution provided by DNA, we show that this negative trend is driven by dinoflagellates while diatoms largely show the opposite trend (Figs. 3.4G and 3.4H). These differences in PDRs align with the classical view of diatoms and dinoflagellates on opposite ends of the r- and K-selected

continuum, while also showing that diatom richness increases under bloom scenarios rather than selecting for a small number of opportunists (Margalef 1978). The results also suggest that a shift to a more stratified less productive regime from climate change may support a more diverse phytoplankton community, but one that has low diatom diversity.

Increases in diatom and dinoflagellate pigments also align with increases in genera that contain harmful bloom-forming taxa (Fig. 3.5 and Supplemental Fig. 3.12). With the harmful diatom genus *Pseudo-nitzschia*, gene expression for toxin biosynthesis was also often elevated under higher fucoxanthin concentrations (Fig. 3.5E). Although increases in these genera does not always imply that there is toxin production, the detection of these pigments, and in particular, the substitution of remotely sensed chlorophyll *a* for phytoplankton pigments will likely improve HAB detection and forecasts. Increases in pigments also corresponds to higher expression of genes specific to their respective group besides with Chlorophytes. These genes fall into diverse metabolic categories, and as connections between gene expression and rate processes emerge as done with quantitative protein measurements (McCain et al. 2024), the detection of pigments may aid the inference of group-specific metabolic activity and support the development of ESMs.

Data availability

Phytoplankton pigment and flow cytometry data are available on CCE-LTER Datazoo: <https://oceaninformatics.ucsd.edu/datazoo/catalogs/ccelter/datasets>. The DNA and RNA sequence data reported in this study have been deposited in the National Center for Biotechnology (NCBI) sequence read archive under the BioProject accession numbers PRJNA555783, PRJNA665326, and PRJNA804265.

Author contributions

RHL, RG, and AEA designed the study approach. RHL, AJR, SMW, AS, and HZ performed the nucleic acid sample collection, processing, and DNA-based analyses. RHL and PV analyzed the metatranscriptomics data. MRL provided the flow cytometry data. RHL, RG, and MRS analyzed the phytoplankton pigment data. RHL, ADB, and AEA synthesized the results. RHL performed the statistical analyses and wrote the manuscript with input from all authors.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank the CCE-LTER scientists, Shonna Dovel and Megan Roadman, and the captains and crews of all research vessels involved. We also thank Mark Ohman and Kathy Barbeau as former and current Lead PIs of the California Current Ecosystem Long-Term Ecological Research (CCE-LTER) site and Brice Semmens as the director of CalCOFI. Zoltán Füssy also assisted in the submission of this manuscript. This work was supported by the National Oceanic and Atmospheric Administration grants NA15OAR4320071 and NA19NOS4780181 (to A.E.A.), the National Science Foundation grants OCE-1026607 (to R.G. and M.R.L.), OCE-1637632 (to R.G. and M.R.L.), OCE-2224726 (to A.E.A. and M.R.S.), and DGE-1650112 (Graduate Research Fellowship to R.H.L.), the Gordon and Betty Moore Foundation grant GBMF3828 (to A.E.A.), and the Simons Foundation Collaboration on

Principles of Microbial Ecosystems (PriME) grant 970820 (to A.E.A.). R.H.L. also received partial support from the LaVerne Noyes Foundation.

Chapter 3, in full, has been submitted for publication and is a reprint of the material as it appears in Lampe, R.H., Rabines, A.J., Wood, S.M., Schulberg, A., Goericke, R., Venepally, P., Zheng, H., Stukel, M.R., Landry, M.R., Barton, A.D., and Allen, A.E. 2024. *EGUsphere* [preprint] <https://doi.org/10.5194/egusphere-2024-3285>. The dissertation author was the primary investigator and author of this paper.

References

- Agusti, S., J. I. González-Gordillo, D. Vaqué, M. Estrada, M. I. Cerezo, G. Salazar, J. M. Gasol, and C. M. Duarte. 2015. Ubiquitous healthy diatoms in the deep sea confirm deep carbon injection by the biological pump. *Nat Comm* **6**: 7608.
- Amaral-Zettler, L. A., E. A. McCliment, H. W. Ducklow, and S. M. Huse. 2009. A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLoS ONE* **4**: e6372.
- Anderson, C. R., R. M. Kudela, M. Kahru, Y. Chao, L. K. Rosenfeld, F. L. Bahr, D. M. Anderson, and T. A. Norris. 2016. Initial skill assessment of the California harmful algae risk mapping (C-HARM) system. *Harmful Algae* **59**: 1-18.
- Anderson, C. R., D. A. Siegel, M. A. Brzezinski, and N. Guillocheau. 2008. Controls on temporal patterns in phytoplankton community structure in the Santa Barbara Channel, California. *J Geophys Res Oceans* **113**: C04038.
- Anderson, D. M., A. D. Cembella, and G. M. Hallegraeff. 2012. Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. *Annu Rev Mar Sci* **4**: 143-176.
- Anderson, D. M., E. Fensin, C. J. Gobler, A. E. Hoeglund, K. A. Hubbard, D. M. Kulis, J. H. Landsberg, K. A. Lefebvre, P. Provoost, and M. L. Richlen. 2021. Marine harmful algal blooms (HABs) in the United States: History, current status and future trends. *Harmful Algae* **102**: 101975.
- Antajan, E., M. J. Chrétiennot-Dinet, C. Leblanc, M. H. Daro, and C. Lancelot. 2004. 19'-hexanoyloxyfucoxanthin may not be the appropriate pigment to trace occurrence and fate of Phaeocystis: the case of *P. globosa* in Belgian coastal waters. *J Sea Res* **52**: 165-177.
- Aramaki, T., R. Blanc-Mathieu, H. Endo, K. Ohkubo, M. Kanehisa, S. Goto, and H. Ogata. 2019. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**: 2251-2252.
- Armstrong, F., C. Stearns, and J. Strickland. 1967. The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment, p. 381-389. *Deep Sea Research and Oceanographic Abstracts*. Elsevier.
- Balch, W. M. 2018. The Ecology, Biogeochemistry, and Optical Properties of Coccolithophores. *Annu Rev Mar Sci* **10**: 71-98.
- Bates, S. S., K. A. Hubbard, N. Lundholm, M. Montresor, and C. P. Leaw. 2018. *Pseudo-nitzschia*, *Nitzschia*, and domoic acid: New research since 2011. *Harmful Algae* **79**: 3-43.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* **57**: 289-300.

Berdjeb, L., A. Parada, D. M. Needham, and J. A. Fuhrman. 2018. Short-term dynamics and interactions of marine protist communities during the spring–summer transition. *ISME J* **12**: 1907-1917.

Bertrand, E. M., J. P. McCrow, A. Moustafa, H. Zheng, J. B. McQuaid, T. O. Delmont, A. F. Post, R. E. Sipler, J. L. Spackeen, K. Xu, D. A. Bronk, D. A. Hutchins, and A. E. Allen. 2015. Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc Natl Acad Sci USA* **112**: 9938-9943.

Bokulich, N. A., B. D. Kaehler, J. R. Rideout, M. Dillon, E. Bolyen, R. Knight, G. A. Huttley, and J. Gregory Caporaso. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**: 90.

Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. Cope, R. Da Silva, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciulek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Priesse, L. B. Rasmussen, A. Rivers, I. I. M. S. Robeson, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, and J. G. Caporaso. 2019. Reproducible, interactive, scalable, and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852-857.

Bowler, C., A. Vardi, and A. E. Allen. 2010. Oceanographic and Biogeochemical Insights from Diatom Genomes. *Annu Rev Mar Sci* **2**: 333-365.

Brunson, J. K., S. M. K. McKinnie, J. R. Chekan, J. P. McCrow, Z. D. Miles, E. M. Bertrand, V. A. Bielinski, H. Luhavaya, M. Oborník, G. J. Smith, D. A. Hutchins, A. E. Allen, and B. S. Moore. 2018. Biosynthesis of the neurotoxin domoic acid in a bloom-forming diatom. *Science* **361**: 1356-1358.

Brunson, J. K., M. Thukral, J. P. Ryan, C. R. Anderson, B. C. Kolody, C. C. James, F. P. Chavez, C. P. Leaw, A. J. Rabines, P. Venepally, Z. Fussy, H. Zheng, R. M. Kudela, G. J. Smith, B. S. Moore, and A. E. Allen. 2024. Molecular forecasting of domoic acid during a pervasive toxic diatom bloom. *Proc Natl Acad Sci USA* **121**: e2319177121.

Catlett, D., D. A. Siegel, P. G. Matson, E. K. Wear, C. A. Carlson, T. S. Lankiewicz, and M. D. Iglesias-Rodríguez. 2023. Integrating phytoplankton pigment and DNA meta-barcoding

observations to determine phytoplankton composition in the coastal ocean. *Limnol Oceanogr* **68**: 361-376.

Cermeño, P., T. Rodríguez-Ramos, M. Dornelas, F. G. Figueiras, E. Marañón, I. G. Teixeira, and S. M. Vallina. 2013. Species richness in marine phytoplankton communities is not correlated to ecosystem productivity. *Mar Ecol Prog Ser* **488**: 1-9.

Cetinić, I., C. S. Rousseaux, I. T. Carroll, A. P. Chase, S. J. Kramer, P. J. Werdell, D. A. Siegel, H. M. Dierssen, D. Catlett, A. Neeley, I. M. Soto Ramos, J. L. Wolny, N. Sadoff, E. Urquhart, T. K. Westberry, D. Stramski, N. Pahlevan, B. N. Seegers, E. Sirk, P. K. Lange, R. A. Vandermeulen, J. R. Graff, J. G. Allen, P. Gaube, L. I. W. McKinna, S. M. McKibben, C. E. Binding, V. S. Calzado, and M. Sayers. 2024. Phytoplankton composition from sPACE: Requirements, opportunities, and challenges. *Remote Sens Environ* **302**: 113964.

Checkley, D. M., and J. A. Barth. 2009. Patterns and processes in the California Current System. *Prog Oceanogr* **83**: 49-64.

Cohen, N. R., H. Alexander, A. I. Krinos, S. K. Hu, and R. H. Lampe. 2022. Marine Microeukaryote Metatranscriptomics: Sample Processing and Bioinformatic Workflow Recommendations for Ecological Applications. *Front Mar Sci* **9**: 867007.

Cotti-Rausch, B. E., M. W. Lomas, E. M. Lachenmyer, E. A. Goldman, D. W. Bell, S. R. Goldberg, and T. L. Richardson. 2016. Mesoscale and sub-mesoscale variability in phytoplankton community composition in the Sargasso Sea. *Deep Sea Res Part I Oceanogr Res Pap* **110**: 106-122.

Cullen, J. J. 2015. Subsurface Chlorophyll Maximum Layers: Enduring Enigma or Mystery Solved? *Annu Rev Mar Sci* **7**: 207-239.

Decelle, J., E. Kayal, E. Bigeard, B. Gallet, J. Bougoure, P. Clode, N. Schieber, R. Templin, E. Hehenberger, G. Prensier, F. Chevalier, Y. Schwab, and L. Guillou. 2022. Intracellular development and impact of a marine eukaryotic parasite on its zombified microalgal host. *ISME J* **16**: 2348-2359.

Decelle, J., S. Romac, R. F. Stern, E. M. Bendif, A. Zingone, S. Audic, M. D. Guiry, L. Guillou, D. Tessier, F. Le Gall, P. Gourvil, A. L. Dos Santos, I. Probert, D. Vaultot, C. de Vargas, and R. Christen. 2015. PhytoREF: a reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy. *Mol Ecol Resour* **15**: 1435-1445.

Deutsch, C., H. Frenzel, J. C. McWilliams, L. Renault, F. Kessouri, E. Howard, J.-H. Liang, D. Bianchi, and S. Yang. 2021. Biogeochemical variability in the California Current System. *Prog Oceanogr* **196**: 102565.

Di Lorenzo, E., and N. Mantua. 2016. Multi-year persistence of the 2014/15 North Pacific marine heatwave. *Nat Clim Change* **6**: 1042-1047.

Ducklow, H., M. Cimino, K. H. Dunton, W. R. Fraser, R. R. Hopcroft, R. Ji, A. J. Miller, M. D. Ohman, and H. M. Sosik. 2022. Marine Pelagic Ecosystem Responses to Climate Variability and Change. *BioScience* **72**: 827-850.

Dupont, C. L., J. P. McCrow, R. Valas, A. Moustafa, N. Walworth, U. Goodenough, R. Roth, S. L. Hogle, J. Bai, Z. I. Johnson, E. Mann, B. Palenik, K. A. Barbeau, J. Craig Venter, and A. E. Allen. 2015. Genomes and gene expression across light and productivity gradients in eastern subtropical Pacific microbial communities. *ISME J* **9**: 1076-1092.

Enright, A. J., S. Van Dongen, and C. A. Ouzounis. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* **30**: 1575-1584.

Eppley, R. W. 1992. Chlorophyll, photosynthesis and new production in the Southern California Bight. *Prog Oceanogr* **30**: 117-150.

Gloor, G. B., J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. 2017. Microbiome datasets are compositional: and this is not optional. *Front Microbiol* **8**: 2224.

Goebel, N., C. Edwards, J. Zehr, M. Follows, and S. Morgan. 2013. Modeled phytoplankton diversity and productivity in the California Current System. *Ecol Model* **264**: 37-47.

Goericke, R. 2011. The structure of marine phytoplankton communities-patterns, rules, and mechanisms. *California Cooperative Oceanic Fisheries Investigations Report* **52**: 182-197.

Gordon, L., J. Jennings, A. Ross, and J. Krest. 1992. A suggested protocol for continuous flow automated analysis of seawater nutrients in the WOCE hydrographic program and the Joint Global Ocean Fluxes Study. *Grp. Tech Rpt*: 92-91.

Guérin, N., M. Ciccarella, E. Flamant, P. Frémont, S. Mangenot, B. Istace, B. Noel, C. Belser, L. Bertrand, K. Labadie, C. Cruaud, S. Romac, C. Bachy, M. Gachenot, E. Pelletier, A. Alberti, O. Jaillon, P. Wincker, J.-M. Aury, and Q. Carradec. 2022. Genomic adaptation of the picoeukaryote *Pelagomonas calceolata* to iron-poor oceans revealed by a chromosome-scale genome sequence. *Commun Biol* **5**: 983.

Guidi, L., S. Chaffron, L. Bittner, D. Eveillard, A. Larhlami, S. Roux, Y. Darzi, S. Audic, L. Berline, J. R. Brum, L. P. Coelho, J. C. I. Espinoza, S. Malviya, S. Sunagawa, C. Dimier, S. Kandels-Lewis, M. Picheral, J. Poulain, S. Searson, L. Stemmann, F. Not, P. Hingamp, S. Speich, M. Follows, L. Karp-Boss, E. Boss, H. Ogata, S. Pesant, J. Weissenbach, P. Wincker, S. G. Acinas, P. Bork, C. de Vargas, D. Iudicone, M. B. Sullivan, J. Raes, E. Karsenti, C. Bowler, G. Gorsky, and C. Tara Oceans Consortium. 2016. Plankton networks driving carbon export in the oligotrophic ocean. *Nature* **532**: 465-470.

Guillou, L., D. Bachar, S. Audic, D. Bass, C. Berney, L. Bittner, C. Boutte, G. Burgaud, C. de Vargas, J. Decelle, J. del Campo, J. R. Dolan, M. Dunthorn, B. Edvardsen, M. Holzmann, W. H. C. F. Kooistra, E. Lara, N. Le Bescot, R. Logares, F. Mahé, R. Massana, M. Montresor, R. Morard, F. Not, J. Pawlowski, I. Probert, A.-L. Sauvadet, R. Siano, T. Stoeck, D. Vaultot, P. Zimmermann, and R. Christen. 2012. The Protist Ribosomal Reference database (PR2): a catalog

of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* **41**: D597-D604.

Guillou, L., M. Viprey, A. Chambouvet, R. M. Welsh, A. R. Kirkham, R. Massana, D. J. Scanlan, and A. Z. Worden. 2008. Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environ Microbiol* **10**: 3349-3365.

Hayward, A., M. H. Pinkerton, and A. Gutierrez-Rodriguez. 2023. phytoclass: A pigment-based chemotaxonomic method to determine the biomass of phytoplankton classes. *Limnol Oceanogr-Meth* **21**: 220-241.

Henriksen, P., B. Riemann, H. Kaas, H. M. Sørensen, and H. L. Sørensen. 2002. Effects of nutrient-limitation and irradiance on marine phytoplankton pigments. *J Plankton Res* **24**: 835-858.

Henson, S. A., B. B. Cael, S. R. Allen, and S. Dutkiewicz. 2021. Future phytoplankton diversity in a changing climate. *Nat Commun* **12**: 5372.

Huang, Y., D. Nicholson, B. Huang, and N. Cassar. 2021. Global Estimates of Marine Gross Primary Production Based on Machine Learning Upscaling of Field Observations. *Glob Biogeochem Cycles* **35**: e2020GB006718.

Ibarbalz, F. M., N. Henry, M. C. Brandão, S. Martini, G. Busseni, H. Byrne, L. P. Coelho, H. Endo, J. M. Gasol, A. C. Gregory, F. Mahé, J. Rigonato, M. Royo-Llloch, G. Salazar, I. Sanz-Sáez, E. Scalco, D. Soviadan, A. A. Zayed, A. Zingone, K. Labadie, J. Ferland, C. Marec, S. Kandels, M. Picheral, C. Dimier, J. Poulain, S. Pisarev, M. Carmichael, S. Pesant, S. G. Acinas, M. Babin, P. Bork, E. Boss, C. Bowler, G. Cochrane, C. de Vargas, M. Follows, G. Gorsky, N. Grimsley, L. Guidi, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels, L. Karp-Boss, E. Karsenti, F. Not, H. Ogata, S. Pesant, N. Poulton, J. Raes, C. Sardet, S. Speich, L. Stemmann, M. B. Sullivan, S. Sunagawa, P. Wincker, M. Babin, E. Boss, D. Iudicone, O. Jaillon, S. G. Acinas, H. Ogata, E. Pelletier, L. Stemmann, M. B. Sullivan, S. Sunagawa, L. Bopp, C. de Vargas, L. Karp-Boss, P. Wincker, F. Lombard, C. Bowler, and L. Zinger. 2019. Global Trends in Marine Plankton Diversity across Kingdoms of Life. *Cell* **179**: 1084-1097.e1021.

Irigoien, X., J. Huisman, and R. P. Harris. 2004. Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature* **429**: 863-867.

Jacox, M. G., M. A. Alexander, N. J. Mantua, J. D. Scott, G. Hervieux, R. S. Webb, and F. E. Werner. 2018. Forcing of multi-year extreme ocean temperatures that impacted California Current living marine resources in 2016. *Bull. Amer. Meteor. Soc* **99**.

James, C. C., A. D. Barton, L. Z. Allen, R. H. Lampe, A. Rabines, A. Schulberg, H. Zheng, R. Goericke, K. D. Goodwin, and A. E. Allen. 2022. Influence of nutrient supply on plankton microbiome biodiversity and distribution in a coastal upwelling region. *Nat Commun* **13**: 2448.

Jeffrey, S. W., S. W. Wright, and M. Zapata. 2011. Microalgal classes and their signature pigments, p. 3-77. *In* S. Roy, C. A. Llewellyn, E. S. Egeland and G. Johnsen [eds.],

Phytoplankton Pigments: Characterization, Chemotaxonomy and Applications in Oceanography. Cambridge Environmental Chemistry Series. Cambridge University Press.

Jones-Kellett, A. E., J. C. McNichol, Y. Raut, K. R. Cain, F. Ribalet, E. V. Armbrust, M. J. Follows, and J. A. Fuhrman. 2024. Amplicon sequencing with internal standards yields accurate picocyanobacteria cell abundances as validated with flow cytometry. *ISME Commun* **4**.

Kahru, M., C. Anderson, A. D. Barton, M. L. Carter, D. Catlett, U. Send, H. M. Sosik, E. L. Weiss, and B. G. Mitchell. 2021. Satellite detection of dinoflagellate blooms off California by UV reflectance ratios. *Elementa: Sci Anthropocene* **9**.

Kamikawa, R., G. Tanifuji, M. Kawachi, H. Miyashita, T. Hashimoto, and Y. Inagaki. 2015. Plastid Genome-Based Phylogeny Pinpointed the Origin of the Green-Colored Plastid in the Dinoflagellate *Lepidodinium chlorophorum*. *Genome Biol Evol* **7**: 1133-1140.

Kanehisa, M., M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353-D361.

Kemp, A. E. S., and T. A. Villareal. 2018. The case of the diatoms and the muddled mandalas: Time to recognize diatom adaptations to stratified waters. *Prog Oceanogr* **167**: 138-149.

Koppelle, S., D. López-Escardó, C. P. D. Brussaard, J. Huisman, C. J. M. Philippart, R. Massana, and S. Wilken. 2022. Mixotrophy in the bloom-forming genus *Phaeocystis* and other haptophytes. *Harmful Algae* **117**: 102292.

Kramer, S. J., L. M. Bolaños, D. Catlett, A. P. Chase, M. J. Behrenfeld, E. S. Boss, E. T. Crockford, S. J. Giovannoni, J. R. Graff, N. Haëntjens, L. Karp-Boss, E. E. Peacock, C. S. Roesler, H. M. Sosik, and D. A. Siegel. 2024a. Toward a synthesis of phytoplankton community composition methods for global-scale application. *Limnol Oceanogr-Meth* **22**: 217-240.

Kramer, S. J., E. L. Jones, M. L. Estapa, N. L. Paul, T. A. Ryneerson, A. E. Santoro, S. Sudek, and C. A. Durkin. 2024b. Ocean carbon export can be predicted from ocean color-based phytoplankton communities. *bioRxiv*: 2024.2009.2021.613760.

Kramer, S. J., and D. A. Siegel. 2019. How Can Phytoplankton Pigments Be Best Used to Characterize Surface Ocean Phytoplankton Groups for Ocean Color Remote Sensing Algorithms? *J Geophys Res Oceans* **124**: 7557-7574.

Kramer, S. J., D. A. Siegel, S. Maritorena, and D. Catlett. 2022. Modeling surface ocean phytoplankton pigments from hyperspectral remote sensing reflectance on global scales. *Remote Sens Environ* **270**: 112879.

Krinos, A. I., M. M. Brisbin, S. K. Hu, N. R. Cohen, T. A. Ryneerson, M. J. Follows, F. Schulz, and H. Alexander. 2023. Missing microbial eukaryotes and misleading meta-omic conclusions. *bioRxiv*: 2023.2007.2030.551153.

Kwiatkowski, L., O. Torres, L. Bopp, O. Aumont, M. Chamberlain, J. R. Christian, J. P. Dunne, M. Gehlen, T. Ilyina, and J. G. John. 2020. Twenty-first century ocean warming, acidification,

deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences* **17**: 3439-3470.

Landry, M. R., S. L. Brown, J. Neveux, C. Dupouy, J. Blanchot, S. Christensen, and R. R. Bidigare. 2003. Phytoplankton growth and microzooplankton grazing in high-nutrient, low-chlorophyll waters of the equatorial Pacific: Community and taxon-specific rate assessments from pigment and flow cytometric analyses. *J Geophys Res Oceans* **108**.

Lewitus, A. J., R. A. Horner, D. A. Caron, E. Garcia-Mendoza, B. M. Hickey, M. Hunter, D. D. Huppert, R. M. Kudela, G. W. Langlois, J. L. Largier, E. J. Lessard, R. RaLonde, J. E. Jack Rensel, P. G. Strutton, V. L. Trainer, and J. F. Tweddle. 2012. Harmful algal blooms along the North American west coast region: History, trends, causes, and impacts. *Harmful Algae* **19**: 133-159.

Li, B., V. Ruotti, R. M. Stewart, J. A. Thomson, and C. N. Dewey. 2009. RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics* **26**: 493-500.

Li, W. K. W. 2002. Macroecological patterns of phytoplankton in the northwestern North Atlantic Ocean. *Nature* **419**: 154-157.

Lim, H. C., S. N. Tan, S. T. Teng, N. Lundholm, E. Orive, H. David, S. Quijano-Scheggia, S. C. Y. Leong, M. Wolf, S. S. Bates, P. T. Lim, and C. P. Leaw. 2018. Phylogeny and species delineation in the marine diatom *Pseudo-nitzschia* (Bacillariophyta) using *cox1*, *LSU*, and *ITS2* rRNA genes: A perspective in character evolution. *J Phycol* **54**: 234-248.

Lin, H., F. I. Kuzminov, J. Park, S. Lee, P. G. Falkowski, and M. Y. Gorbunov. 2016. The fate of photons absorbed by phytoplankton in the global ocean. *Science* **351**: 264-267.

Lin, Y., S. Gifford, H. Ducklow, O. Schofield, and N. Cassar. 2019. Towards Quantitative Microbiome Community Profiling Using Internal Standards. *Appl Environ Microbiol* **85**: e02634-02618.

Lopes dos Santos, A., C. G rikas Ribeiro, D. Ong, L. Garczarek, X. L. Shi, S. D. Nodder, D. Vaultot, and A. Guti rrez-Rodr guez. 2022. Chapter 11 - Phytoplankton diversity and ecology through the lens of high throughput sequencing technologies, p. 353-413. *In* L. A. Clementson, R. S. Eriksen and A. Willis [eds.], *Advances in Phytoplankton Ecology*. Elsevier.

Lopes dos Santos, A., P. Gourvil, M. Tragin, M.-H. No l, J. Decelle, S. Romac, and D. Vaultot. 2017. Diversity and oceanic distribution of prasinophytes clade VII, the dominant group of green algae in oceanic waters. *The ISME Journal* **11**: 512-528.

Lund, S. 2024. California Current temperatures adjacent to Southern California: 1949–2020. *Quaternary International* **705**: 104-108.

Ma, Z. S. 2018. Measuring microbiome diversity and similarity with Hill numbers, p. 157-178. *Metagenomics*. Elsevier.

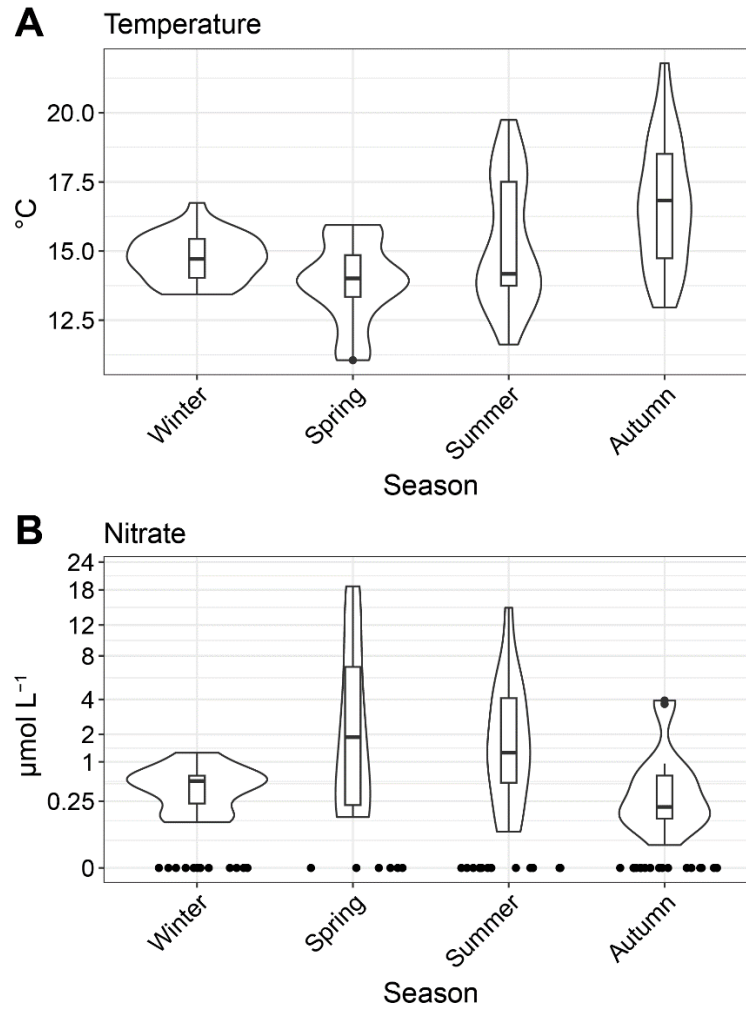
- Mackey, M., D. Mackey, H. Higgins, and S. Wright. 1996. CHEMTAX-a program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton. *Mar Ecol Prog Ser* **144**: 265-283.
- Margalef, R. 1978. Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanol acta* **1**: 493-509.
- Martin, J. L., I. Santi, P. Pitta, U. John, and N. Gypens. 2022. Towards quantitative metabarcoding of eukaryotic plankton: an approach to improve 18S rRNA gene copy number bias. *Metabarcoding Metagenomics* **6**.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 10-12.
- McCain, J. S. P., G. L. Britten, S. R. Hackett, M. J. Follows, and G.-W. Li. 2024. Microbial reaction rate estimation using proteins and proteomes. *bioRxiv*: 2024.2008.2013.607198.
- Mendes, C. R. B., V. M. Tavano, T. S. Dotto, R. Kerr, M. S. De Souza, C. A. E. Garcia, and E. R. Secchi. 2018. New insights on the dominance of cryptophytes in Antarctic coastal waters: a case study in Gerlache Strait. *Deep Sea Research Part II: Topical Studies in Oceanography* **149**: 161-170.
- Monger, B. C., and M. R. Landry. 1993. Flow Cytometric Analysis of Marine Bacteria with Hoechst 33342. *Appl Environ Microbiol* **59**: 905-911.
- Moran, M. A., B. Satinsky, S. M. Gifford, H. Luo, A. Rivers, L.-K. Chan, J. Meng, B. P. Durham, C. Shen, V. A. Varaljay, C. B. Smith, P. L. Yager, and B. M. Hopkinson. 2012. Sizing up metatranscriptomics. *ISME J* **7**: 237-243.
- Moreno, A. R., C. Anderson, R. M. Kudela, M. Sutula, C. Edwards, and D. Bianchi. 2022. Development, calibration, and evaluation of a model of *Pseudo-nitzschia* and domoic acid production for regional ocean modeling studies. *Harmful Algae* **118**: 102296.
- Nayar, S., and L. M. Chou. 2003. Relative efficiencies of different filters in retaining phytoplankton for pigment and productivity studies. *Estuar Coast Shelf Sci* **58**: 241-248.
- Neeley, A. R., M. W. Lomas, A. Mannino, C. Thomas, and R. Vandermeulen. 2022. Impact of Growth Phase, Pigment Adaptation, and Climate Change Conditions on the Cellular Pigment and Carbon Content of Fifty-One Phytoplankton Isolates. *J Phycol* **58**: 669-690.
- Not, F., R. Siano, W. H. C. F. Kooistra, N. Simon, D. Vaultot, and I. Probert. 2012. Chapter One - Diversity and Ecology of Eukaryotic Marine Phytoplankton, p. 1-53. *In* G. Piganeau [ed.], *Advances in Botanical Research*. Academic Press.
- Novák Vanclová, A. M., and R. G. Dorrell. 2024. Complex plastids across the eukaryotes: an overview of inherited and convergently evolved characters. *Endosymbiotic Organelle Acquisition: Solutions to the Problem of Protein Localization and Membrane Passage*: 39-88.

- Parada, A. E., D. M. Needham, and J. A. Fuhrman. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* **18**: 1403-1414.
- Partensky, F., W. R. Hess, and D. Vaultot. 1999. *Prochlorococcus*, a Marine Photosynthetic Prokaryote of Global Significance. *Microbiol Mol Biol Rev* **63**: 106-127.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg. 2011. Scikit-learn: Machine learning in Python. *J Mach Learn Res* **12**: 2825-2830.
- Pierella Karlusich, J. J., F. M. Ibarbalz, and C. Bowler. 2020. Phytoplankton in the Tara Ocean. *Annu Rev Mar Sci* **12**: 233-265.
- Pierella Karlusich, J. J., E. Pelletier, L. Zinger, F. Lombard, A. Zingone, S. Colin, J. M. Gasol, R. G. Dorrell, N. Henry, E. Scalco, S. G. Acinas, P. Wincker, C. de Vargas, and C. Bowler. 2023. A robust approach to estimate relative phytoplankton cell abundances from metagenomes. *Mol Ecol Resour* **23**: 16-40.
- Podell, S., and T. Gaasterland. 2007. DarkHorse: a method for genome-wide prediction of horizontal gene transfer. *Genome Biol* **8**: R16.
- Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188-7196.
- Raven, J. 1998. The twelfth Tansley Lecture. Small is beautiful: the picophytoplankton. *Funct Ecol* **12**: 503-513.
- Rho, M., H. Tang, and Y. Ye. 2010. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* **38**: e191.
- Richardson, T. L. 2022. The colorful world of cryptophyte phycobiliproteins. *J Plankton Res* **44**: 806-818.
- Roberts, M. E., M. P. Bhatia, E. Rowland, P. L. White, S. Waterman, M. A. Cavaco, P. Williams, J. N. Young, J. S. Spence, J.-É. Tremblay, J.-C. Montero-Serrano, and E. M. Bertrand. 2024. Rubisco in high Arctic tidewater glacier-marine systems: A new window into phytoplankton dynamics. *Limnol Oceanogr* **69**: 802-817.
- Rost, B., and U. Riebesell. 2004. Coccolithophores and the biological pump: responses to environmental changes, p. 99-125. *Coccolithophores: from molecular processes to global impact*. Springer.
- Ryan-Keogh, T. J., S. J. Thomalla, P. M. S. Monteiro, and A. Tagliabue. 2023. Multidecadal trend of increasing iron stress in Southern Ocean phytoplankton. *Science* **379**: 834-840.

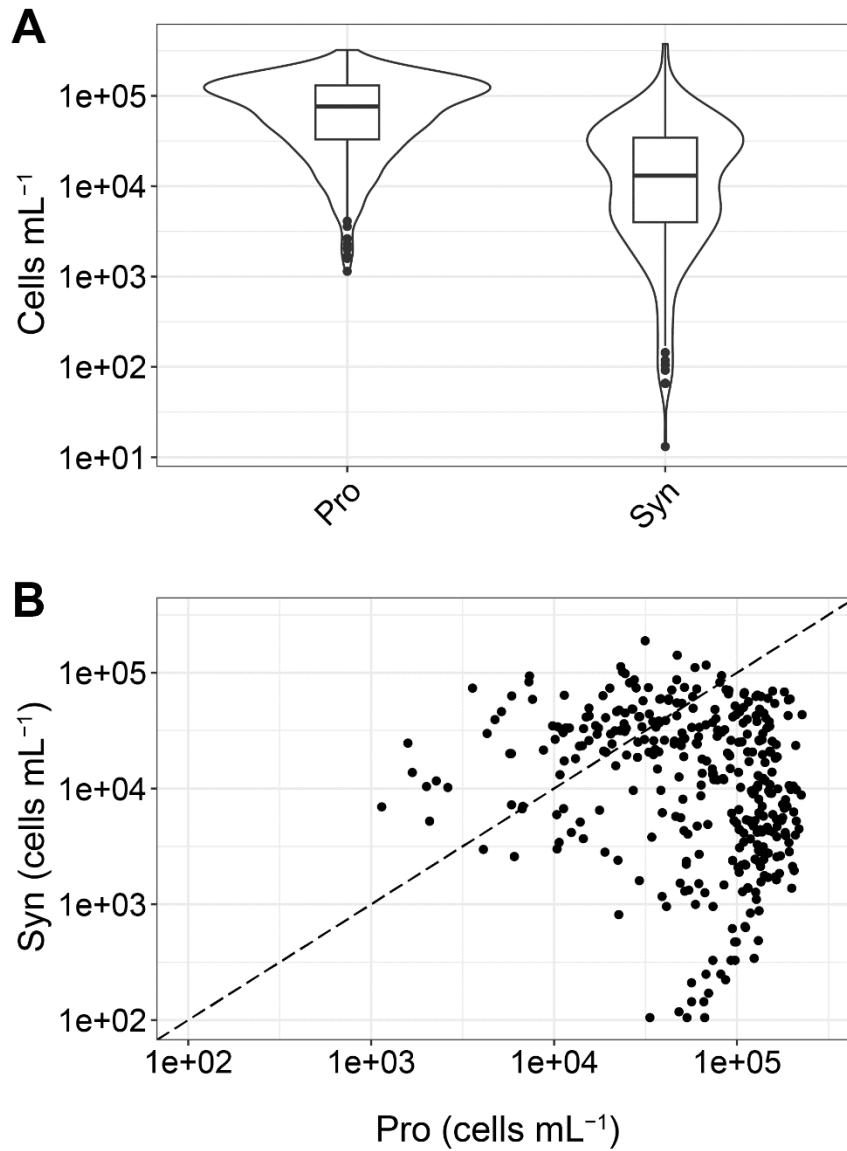
- Saito, M. A., H. Alexander, H. M. Benway, P. W. Boyd, M. Gledhill, E. B. Kujawinski, N. M. Levine, M. Maheigan, A. Marchetti, I. Obernosterer, A. E. Santoro, D. Shi, K. Suzuki, A. Tagliabue, B. S. Twining, and M. T. Maldonado. 2024. The Dawn of the BioGeoSCAPES Program: Ocean Metabolism and Nutrient Cycles on a Changing Planet. *Oceanography* **37**: 162-166.
- Saito, M. A., M. R. McIlvin, D. M. Moran, A. E. Santoro, C. L. Dupont, P. A. Rafter, J. K. Saunders, D. Kaul, C. H. Lamborg, and M. Westley. 2020. Abundant nitrite-oxidizing metalloenzymes in the mesopelagic zone of the tropical Pacific Ocean. *Nat Geosci* **13**: 355-362.
- Sathyendranath, S., R. J. W. Brewin, S. Ciavatta, T. Jackson, G. Kulk, B. Jönsson, V. M. Vicente, and T. Platt. 2023. Ocean Biology Studied from Space. *Surveys in Geophysics* **44**: 1287-1308.
- Satterthwaite, E. V., A. E. Allen, R. H. Lampe, Z. Gold, A. R. Thompson, N. Bowlin, R. Swalethorp, K. D. Goodwin, E. L. Hazen, and S. J. Bograd. 2023. Toward identifying the critical ecological habitat of larval fishes. *Oceanography* **36**: 90-93.
- Schmieder, R., Y. W. Lim, and R. Edwards. 2011. Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* **28**: 433-435.
- Smith Jr, W. O., and S. Trimborn. 2023. Phaeocystis: A global enigma. *Annu Rev Mar Sci* **16**.
- Smith, V. H. 2007. Microbial diversity–productivity relationships in aquatic ecosystems. *FEMS Microbiol Ecol* **62**: 181-186.
- Stoecker, D. K., P. J. Hansen, D. A. Caron, and A. Mitra. 2017. Mixotrophy in the marine plankton. *Annu Rev Mar Sci* **9**: 311-335.
- Strzepek, R. F., B. L. Nunn, L. T. Bach, J. A. Berges, E. B. Young, and P. W. Boyd. 2022. The ongoing need for rates: can physiology and omics come together to co-design the measurements needed to understand complex ocean biogeochemistry? *J Plankton Res* **44**: 485-495.
- Tagliabue, A. 2023. ‘Oceans are hugely complex’: modelling marine microbes is key to climate forecasts. *Nature* **623**: 250-252.
- Ternon, E., M. L. Carter, L. Cancelada, R. H. Lampe, A. E. Allen, C. R. Anderson, K. A. Prather, and W. H. Gerwick. 2023. Yessotoxin production and aerosolization during the unprecedented red tide of 2020 in southern California. *Elementa: Sci Anthropocene* **11**.
- Ting, C. S., C. Hsieh, S. Sundararaman, C. Mannella, and M. Marko. 2007. Cryo-Electron Tomography Reveals the Comparative Three-Dimensional Architecture of *Prochlorococcus*, a Globally Important Marine Cyanobacterium. *J Bacteriol* **189**: 4485-4493.
- Trainer, V. L., G. Pitcher, B. Reguera, and T. Smayda. 2010. The distribution and impacts of harmful algal bloom species in eastern boundary upwelling systems. *Prog Oceanogr* **85**: 33-52.

- Vallina, S. M., M. J. Follows, S. Dutkiewicz, J. M. Montoya, P. Cermenó, and M. Loreau. 2014. Global relationship between phytoplankton diversity and productivity in the ocean. *Nat Commun* **5**: 4299.
- Venrick, E. 1998. Spring in the California Current: the distribution of phytoplankton species, April 1993 and April 1995. *Mar Ecol Prog Ser* **167**: 73-88.
- Venrick, E. 2002. Floral patterns in the California Current System off southern California: 1990-1996. *J Mar Res* **60**: 171-189.
- Venrick, E. L. 2012. Phytoplankton in the California Current system off southern California: Changes in a changing environment. *Prog Oceanogr* **104**: 46-58.
- Wang, J.-X., F.-Z. Kong, H.-X. Geng, Y. Zhao, W.-B. Guan, C. He, Z.-J. Kang, W. Guo, Z.-X. Zhou, Q.-C. Zhang, and R.-C. Yu. 2022. Pigment Characterization of the Giant-Colony-Forming Haptophyte *Phaeocystis globosa* in the Beibu Gulf Reveals Blooms of Different Origins. *Appl Environ Microbiol* **88**: e01654-01621.
- Wood, S. N. 2017. Generalized additive models: an introduction with R. Chapman and Hall/CRC.
- Worden, A. Z., J. Janouskovec, D. McRose, A. Engman, R. M. Welsh, S. Malfatti, S. G. Tringe, and P. J. Keeling. 2012. Global distribution of a wild alga revealed by targeted metagenomics. *Curr Biol* **22**: R675-R677.
- Yeh, Y.-C., J. McNichol, D. M. Needham, E. B. Fichot, L. Berdjeb, and J. A. Fuhrman. 2021. Comprehensive single-PCR 16S and 18S rRNA community analysis validated with mock communities, and estimation of sequencing bias against 18S. *Environ Microbiol* **23**: 3240-3250.
- Zapata, M., S. Fraga, F. Rodríguez, and J. L. Garrido. 2012. Pigment-based chloroplast types in dinoflagellates. *Mar Ecol Prog Ser* **465**: 33-52.
- Zapata, M., S. W. Jeffrey, S. W. Wright, F. Rodríguez, J. L. Garrido, and L. Clementson. 2004. Photosynthetic pigments in 37 species (65 strains) of Haptophyta: implications for oceanography and chemotaxonomy. *Mar Ecol Prog Ser* **270**: 83-102.
- Zheng, B., A. J. Lucas, P. J. S. Franks, T. L. Schlosser, C. R. Anderson, U. Send, K. Davis, A. D. Barton, and H. M. Sosik. 2023. Dinoflagellate vertical migration fuels an intense red tide. *Proc Natl Acad Sci USA* **120**: e2304590120.
- Zhu, F., R. Massana, F. Not, D. Marie, and D. Vaultot. 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol* **52**: 79-92.

CHAPTER 3 SUPPLEMENTAL INFORMATION

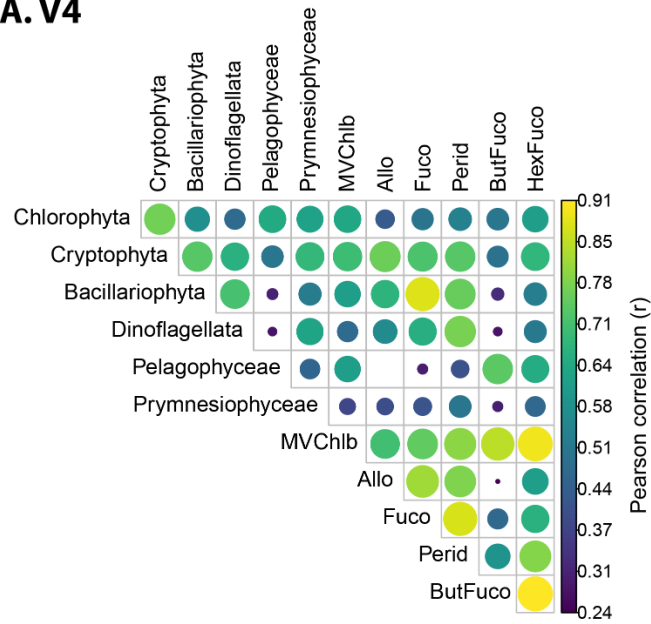


Supplemental Figure 3.1: (A) Temperature (°C) and (B) nitrate concentrations (μmol L⁻¹) among seasons for the RNA samples.

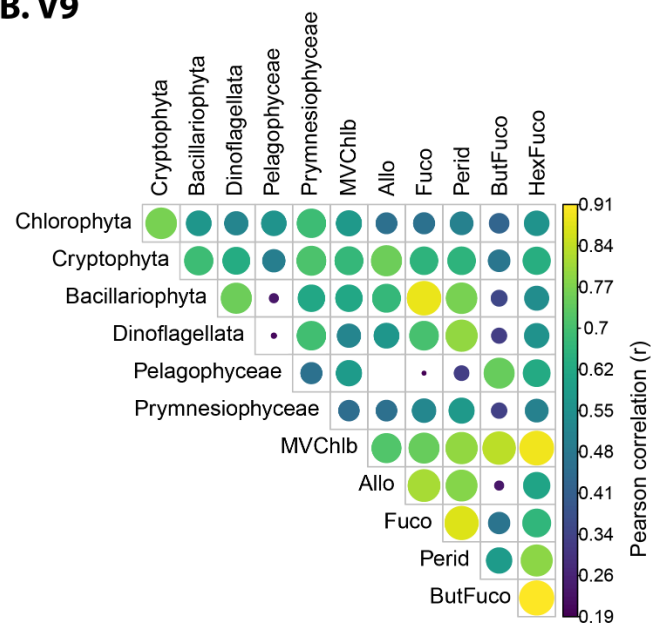


Supplemental Figure 3.2: Cell abundances (cells mL^{-1}) for *Prochlorococcus* (Pro) and *Synechococcus* (Syn) as (A) distributions and (B) sample-by-sample comparisons where the dashed line represents a one-to-one relationship.

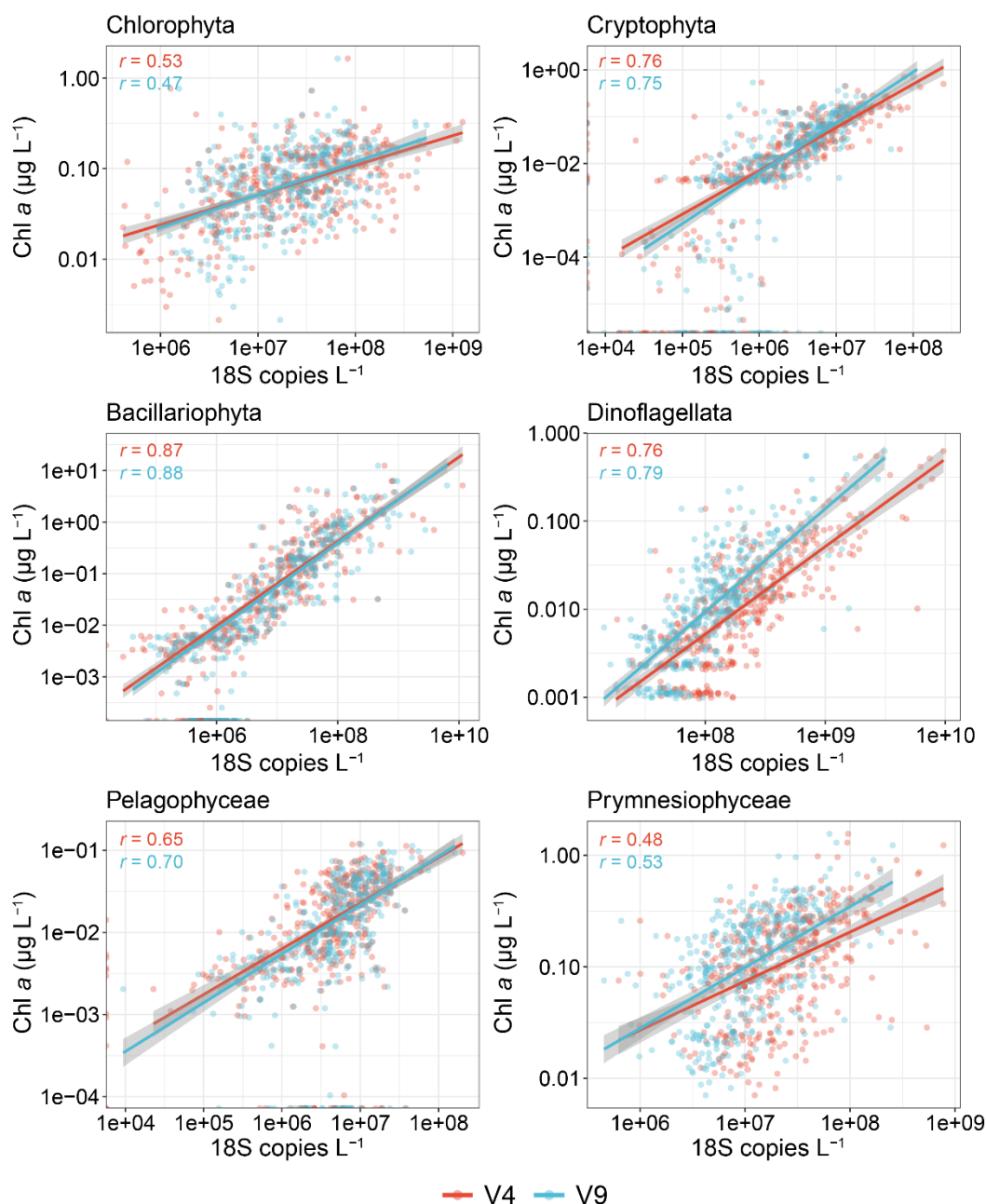
A. V4



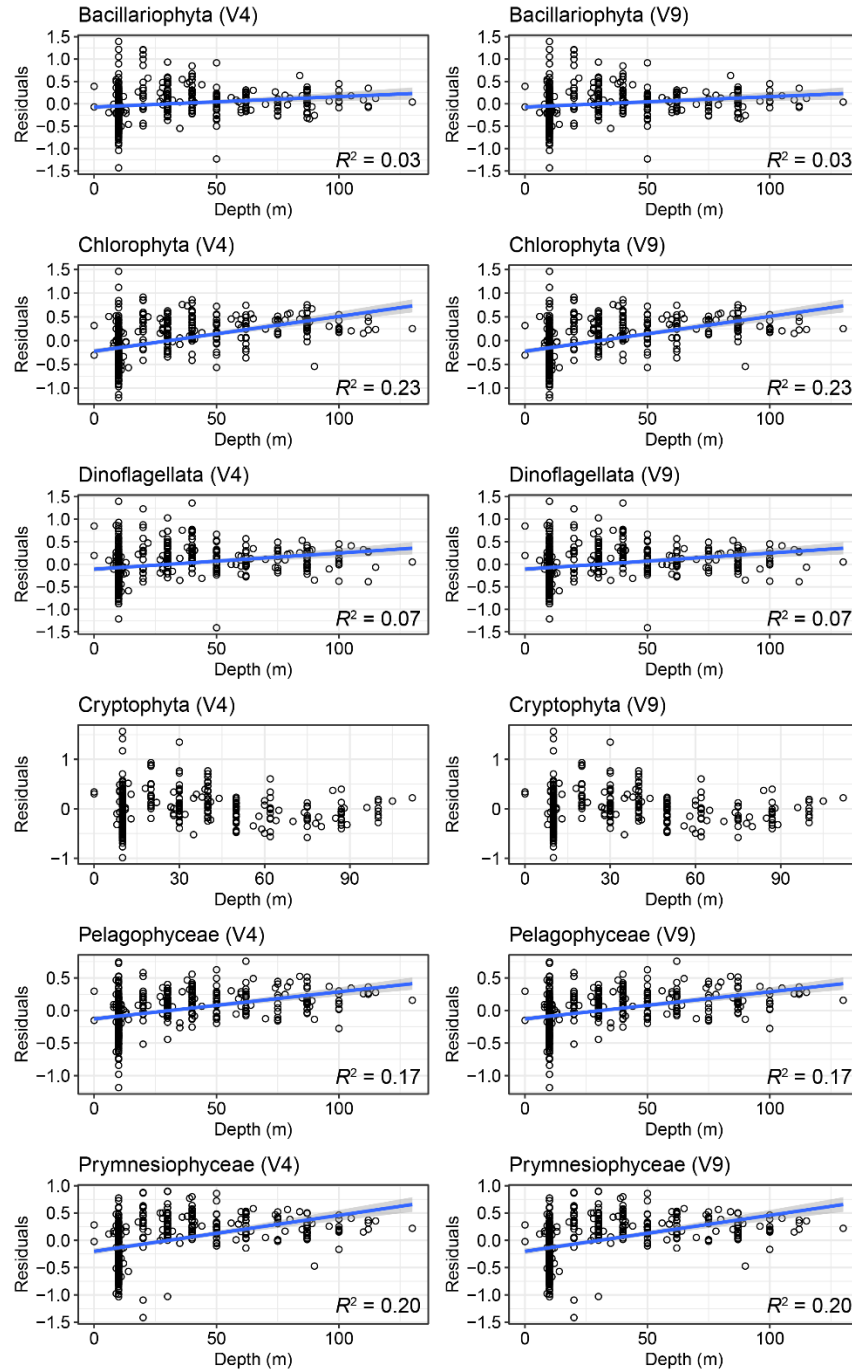
B. V9



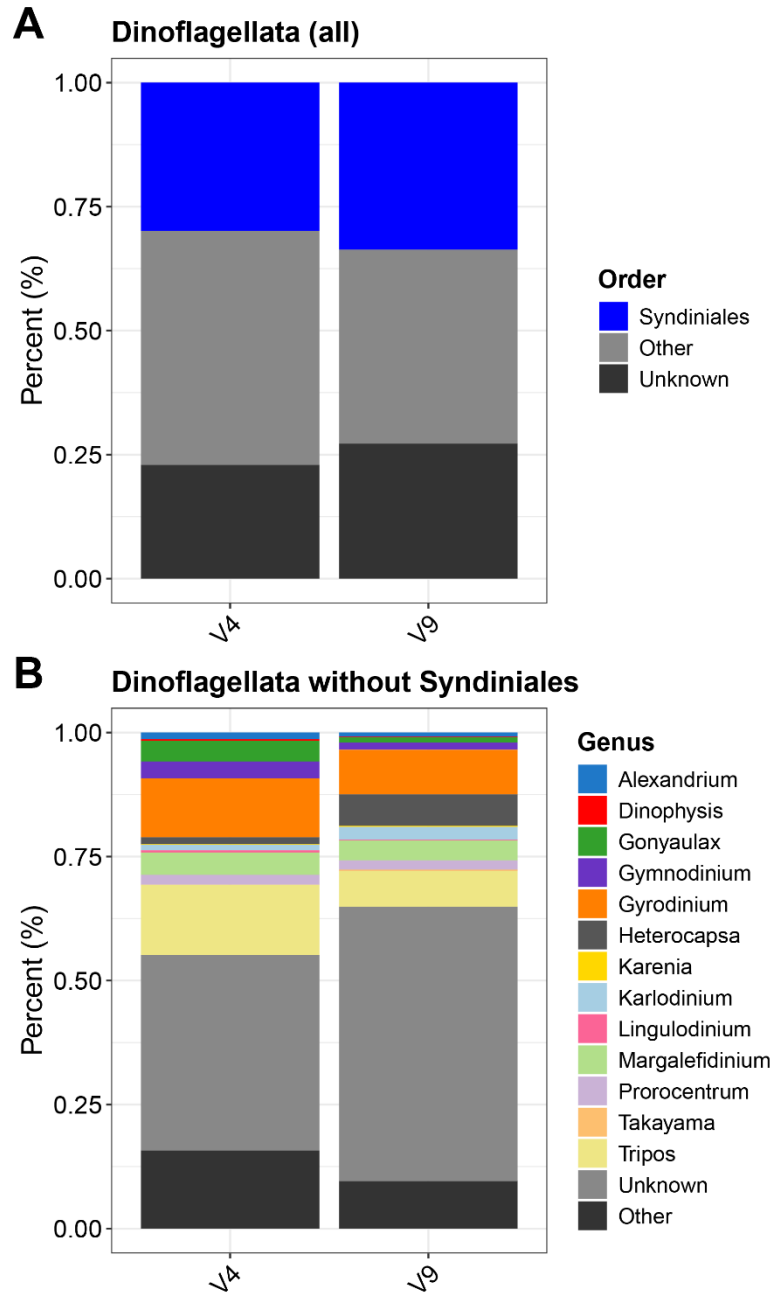
Supplemental Figure 3.3: Pearson correlations among absolute abundances of phytoplankton groups and diagnostic pigments for (A) 18S-V4 and (B) 18S-V9 data. Correlations are not shown where $P > 0.05$.



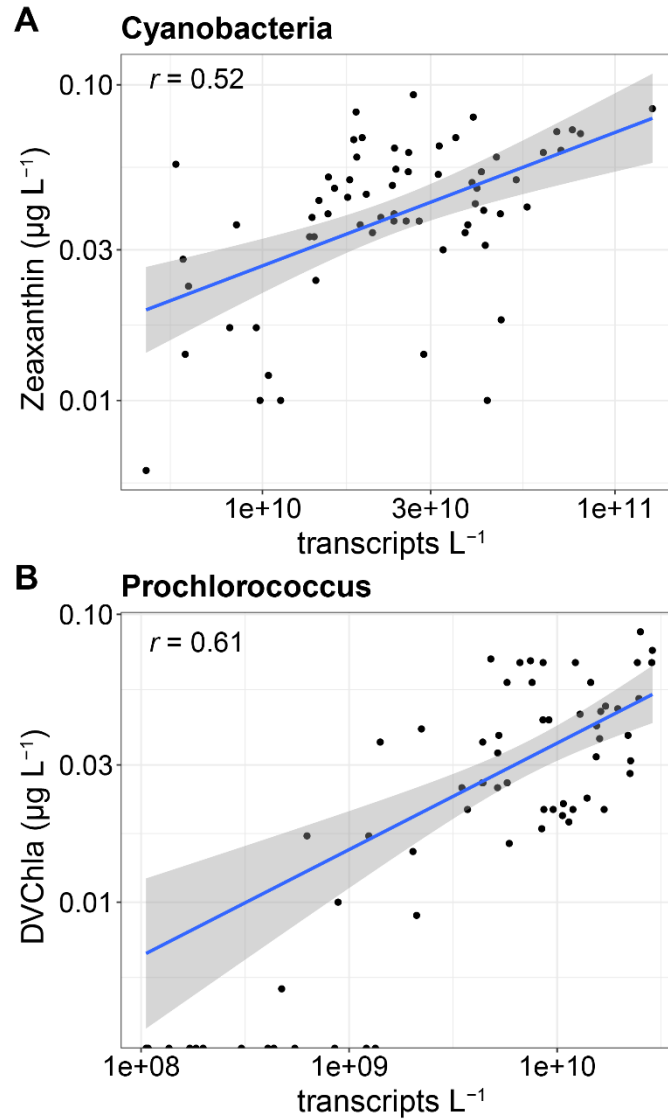
Supplemental Figure 3.4: Correlations between 18S-V4 (red) or 18S-V9 (blue) absolute abundances (copies L⁻¹) and taxon-specific contributions to total chlorophyll estimated with phytoclass (µg L⁻¹) for each major eukaryotic phytoplankton group. Pearson correlations are significant ($P < 0.05$) and coefficients are displayed in each panel. Lines with 95% confidence intervals shows linear models with significant relationships between variables ($P < 0.05$).



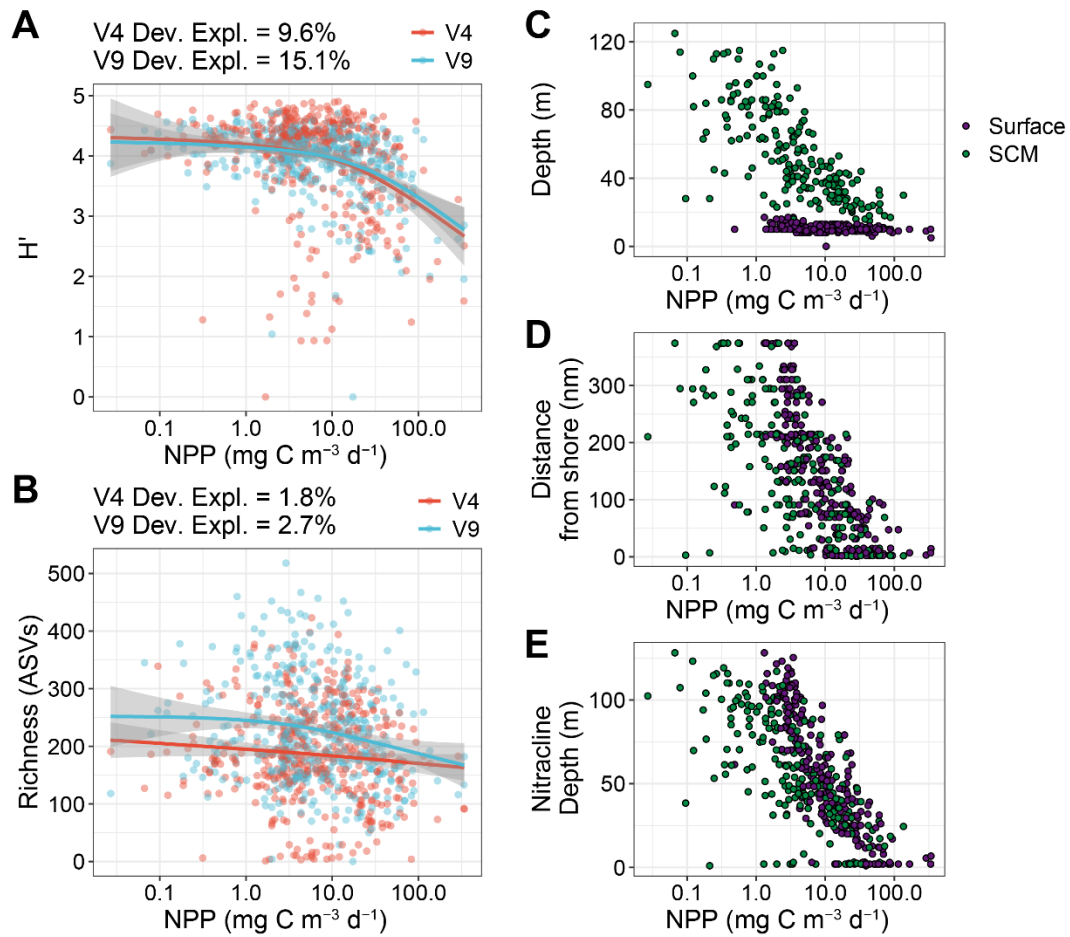
Supplemental Figure 3.5: Linear regressions for depth (m) in predicting residuals from linear models on the absolute abundance and diagnostic pigments shown in Fig. 2B. Coefficients of determination (R^2) and regression lines are shown when $P < 0.05$.



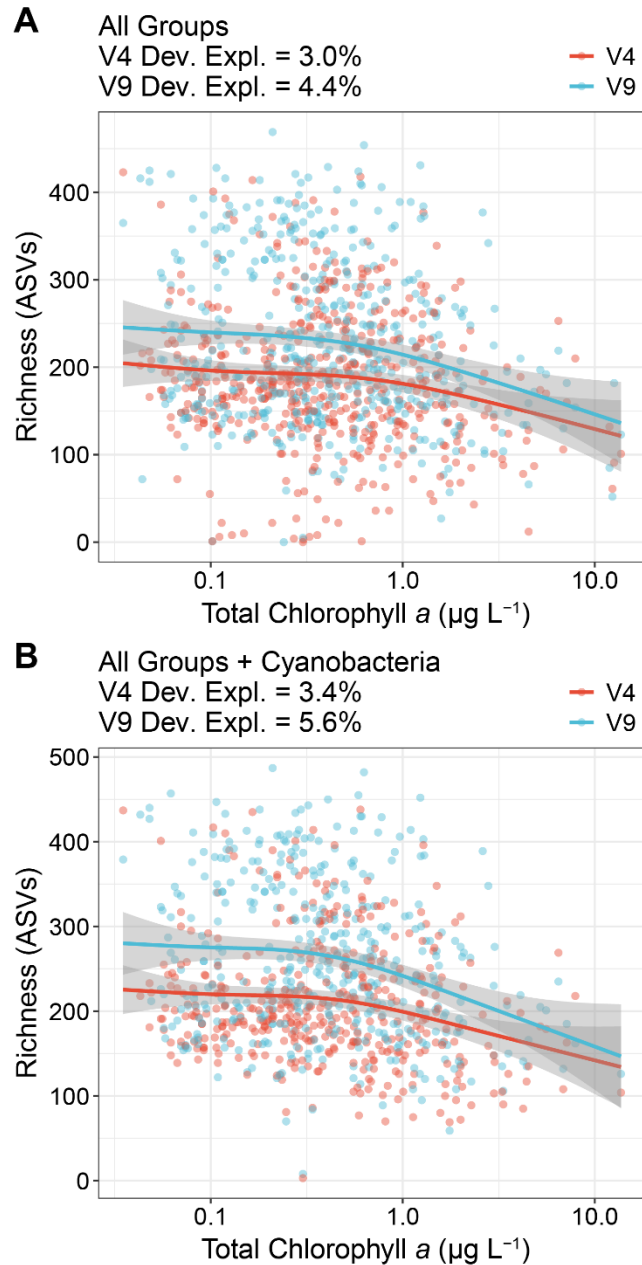
Supplemental Figure 3.6: Relative abundances based on total copies L^{-1} of Dinoflagellata ASVs on the (A) order level including Syndiniales and (B) genus level when Syndiniales are removed.



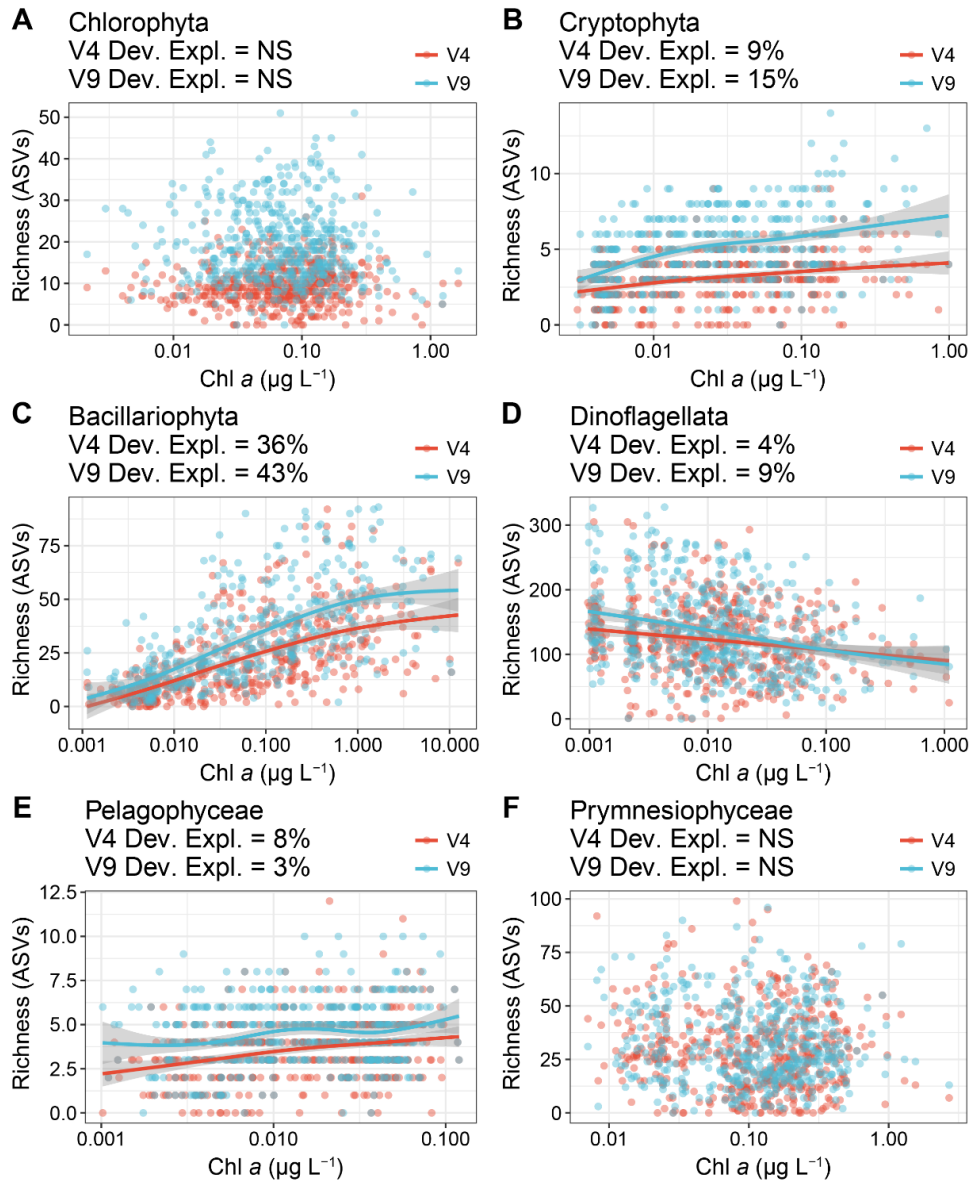
Supplemental Figure 3.7: Correlations between diagnostic pigment concentrations and the absolute abundances of mRNA (transcripts L^{-1}) for (A) all cyanobacteria and zeaxanthin and (B) Prochlorococcus and divinyl chlorophyll a (DVChla). Only samples from the upper 20 m are included. Pearson correlations are significant ($P < 0.05$) and coefficients are displayed in each panel. Lines with 95% confidence intervals show linear models with significant relationships between variables ($P < 0.05$).



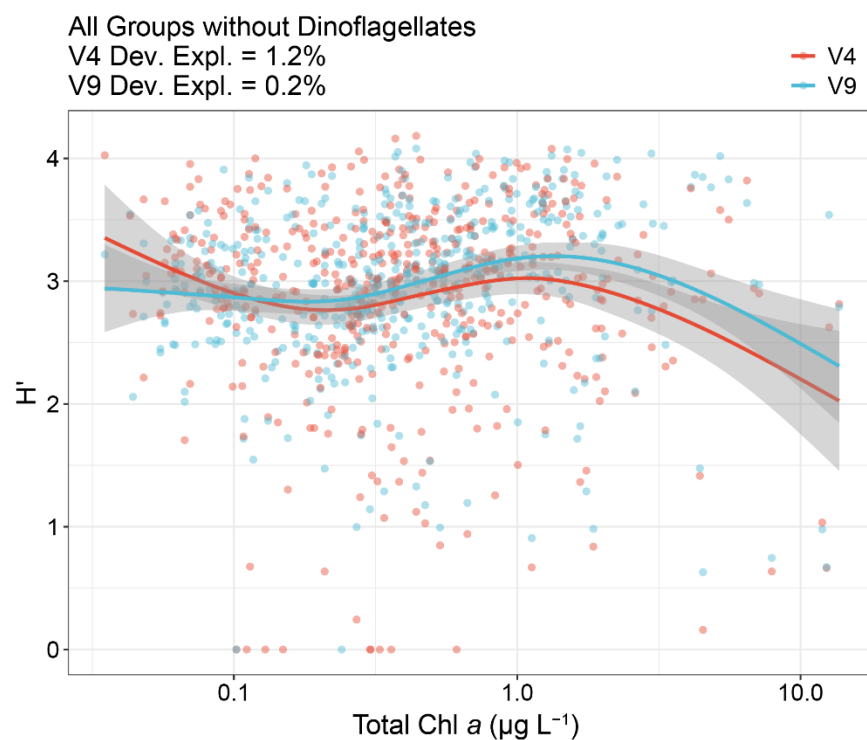
Supplemental Figure 3.8: Productivity-diversity relationships in the region with net primary productivity (NPP) and both (a) the Shannon Index and (b) richness (ASVs) using both the 18S-V4 (red) and 18S-V9 (blue) data. Lines represent GAMs and corresponding 95% confidence intervals where significant ($P < 0.05$). The deviance explained by each GAM is shown above each panel for each amplicon. (C-E) Environmental variables against total chlorophyll a concentrations coloured by near-surface (purple) or subsurface chlorophyll max (SCM, green) samples.



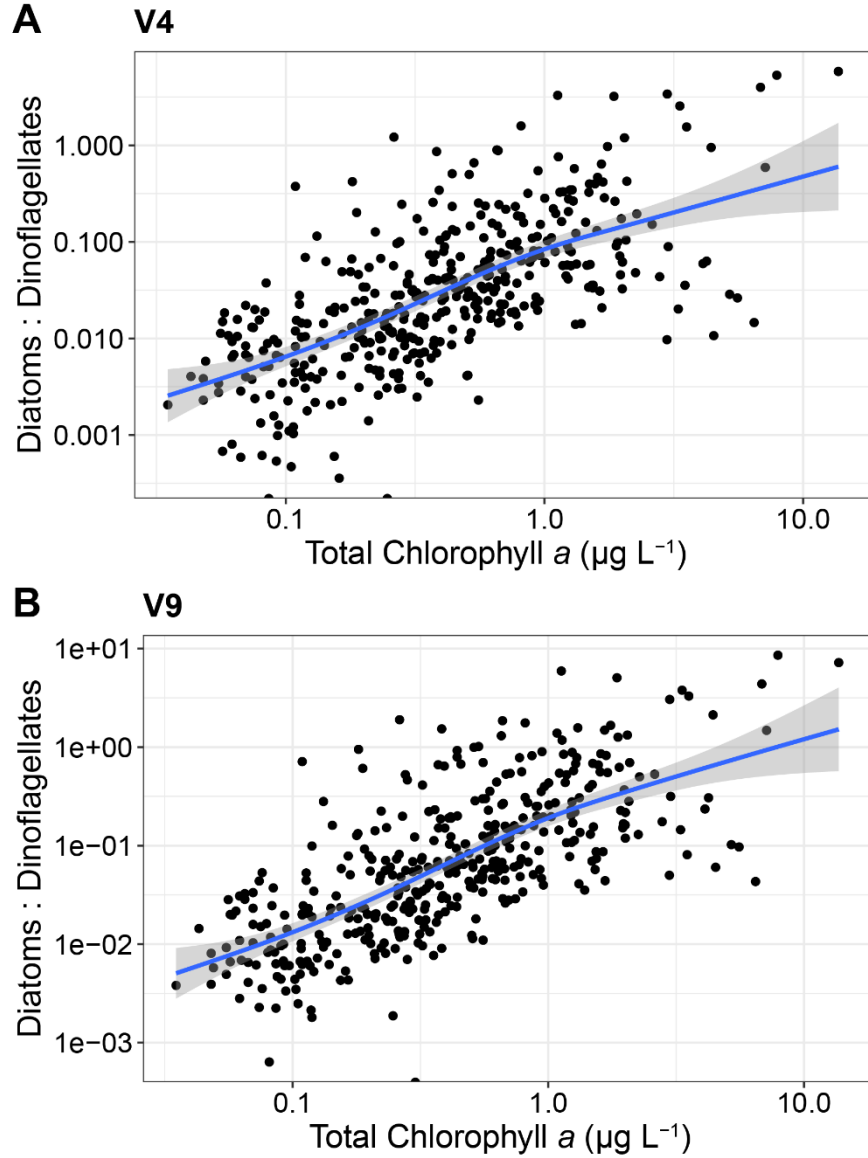
Supplemental Figure 3.9: Productivity-diversity relationship in the region for (A) all eukaryotic phytoplankton groups or (B) all eukaryotic groups and cyanobacteria with biomass (chlorophyll a concentrations) as a proxy for productivity and diversity expressed as the richness (number of ASVs) for both the 18S-V4 (red) and 18S-V9 (blue) data. Lines represent significant GAMs and corresponding 95% confidence intervals ($P < 0.05$). The deviance explained by each GAM is shown for each panel for each amplicon.



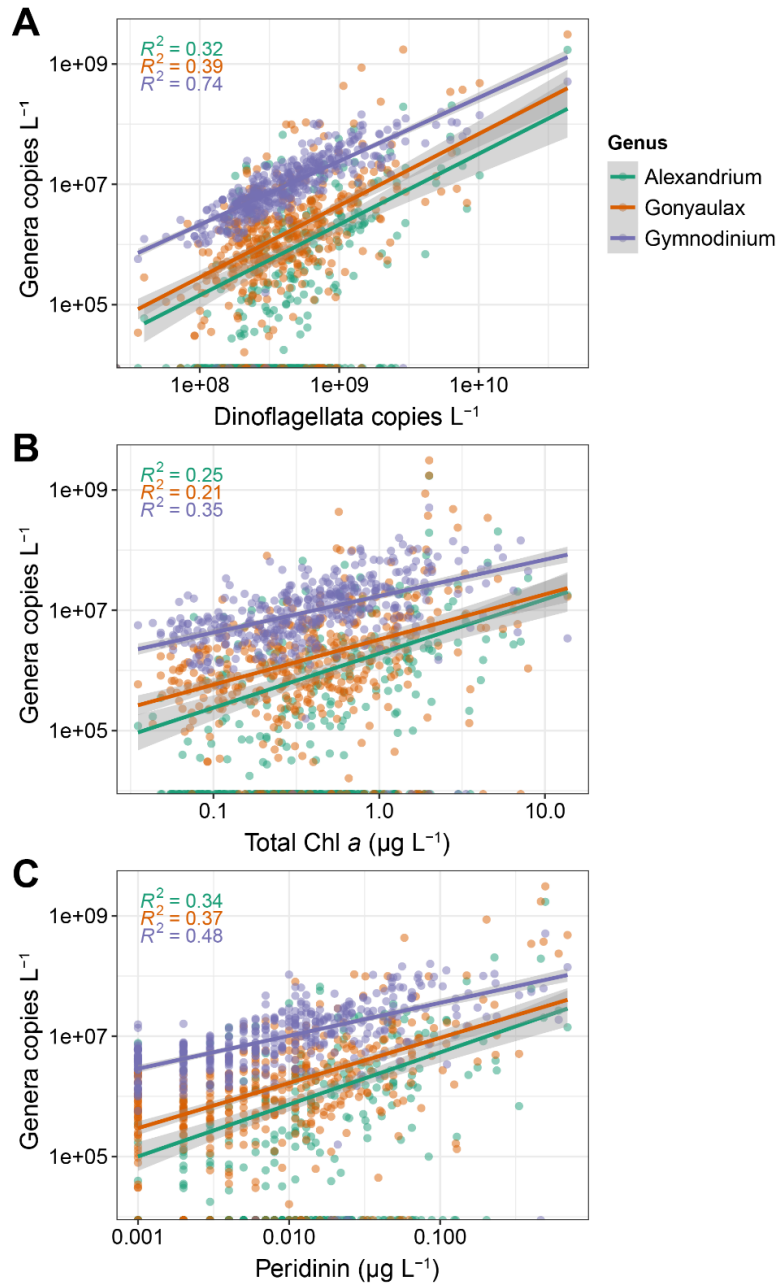
Supplemental Figure 3.10: Productivity-diversity relationships for each individual phytoplankton groups using richness (ASVs) against taxon-specific biomass estimated with phytoclass. Lines represent GAMs and corresponding 95% confidence intervals where significant ($P < 0.05$). The deviance explained by each GAM or “NS” for not significant ($P > 0.05$) is shown above each panel for each amplicon.



Supplemental Figure 3.11: Productivity-diversity relationships in the region for all eukaryotic phytoplankton groups except for dinoflagellates with biomass (chlorophyll a concentrations) as a proxy for productivity and diversity expressed as the Shannon Index (H') for both the 18S-V4 (red) and 18S-V9 (blue) data. Lines represent significant GAMs and corresponding 95% confidence intervals ($P < 0.05$). The deviance explained by each GAM is shown above the plot for each amplicon.



Supplemental Figure 3.12: Diatom-to-dinoflagellate ratio for the (A) 18S-V4 and (B) 18S-V9 data against total chlorophyll *a* concentrations. The line and 95% confidence interval represent the GAM ($P < 0.0001$) that explains 15.0% (V4) or 18.2% (V9) of the deviance.



Supplemental Figure 3.13: Dinoflagellate genus-level abundances against (A) total dinoflagellate copies L^{-1} , (B) total chlorophyll a concentrations, and (C) peridinin concentrations.

Supplemental Table 3.1: Pearson correlation coefficients for dinoflagellates with different groups removed and peridinin.

Removed Group	Amplicon			
	Relative		Absolute	
	V4	V9	V4	V9
None	0.18	0.20	0.73	0.65
Syndiniales	0.42	0.35	0.77	0.70
Kareniceae	0.18	0.20	0.73	0.65
Dinophysis	0.17	0.20	0.73	0.65

CHAPTER 4 - COASTAL UPWELLING DYNAMICS GOVERN DISTINCT PATTERNS IN DIATOM ABUNDANCES, DIVERSITY, AND GENE EXPRESSION

Robert H. Lampe, Ariel J. Rabines, Anne Schulberg, Ralf Goericke, Pratap Venepally, Michael Sorenson, Hong Zheng, Brice X. Semmens, Kelly D. Goodwin, Andrew D. Barton, Andrew E. Allen

Abstract

Coastal upwelling regions, such as the California Current Ecosystem (CCE), are among the most biologically productive areas in the ocean. In these regions, seasonal upwelling fuels intense phytoplankton blooms dominated by diatoms, which in turn support biogeochemical cycling, carbon sequestration, and productive fisheries. Here we contrast pelagic diatom abundances, diversity, and gene expression with other major phytoplankton groups in the CCE via standardized oceanographic surveys that capture extensive spatiotemporal and biogeochemical variability within hundreds of samples for 18S metabarcoding and metatranscriptomics. We show that in contrast to other phytoplankton groups, seasonal upwelling conditions are strongly associated with increases in diatom abundances and diversity as well as differences in gene expression. Overall, the community is dominated by the bloom-forming genera *Chaetoceros*, *Thalassiosira*, and *Pseudo-nitzschia*; however, members of these genera are dispersal limited, even over relatively small spatial scales. Rather than becoming established offshore, these diatoms appear to sink out of the euphotic zone and contribute to carbon export. Furthermore, by integrating these data with others, we estimate global pelagic diatom diversity with the V9 region of the 18S rRNA gene to be 11,432 ASVs.

Introduction

Diatoms (Bacillariophyta) are among the most diverse and prolific phytoplankton groups in the ocean (Armbrust, 2009; Bowler et al., 2010). In particular, they frequently dominate phytoplankton blooms in coastal upwelling regions associated with eastern boundary currents such as the California Current Ecosystem (CCE) (Allen, 1945; Estrada and Blasco, 1985; Goericke, 2011; Taylor and Landry, 2018; Venrick, 2012). These regions are highly dynamic, where in certain areas, seasonal upwelling-favorable winds cause cold, nutrient-rich water to be delivered to the euphotic zone, fueling phytoplankton growth and primary productivity (Chavez and Messié, 2009). As these diatom-dominated phytoplankton blooms lead to some of the highest productivity in the ocean, diatoms in upwelling regions disproportionally contribute to global biogeochemical cycling and the support of higher trophic levels (Capone and Hutchins, 2013).

Despite their significant presence and ecological importance, the extent of diatom abundances and diversity within the spatiotemporal and biogeochemical variability of upwelling ecosystems remains relatively unexplored. Indeed, diatoms often bloom during upwelling events and regularly account for a majority of phytoplankton biomass as observed from phytoplankton pigments, although pigments lack deeper taxonomic resolution (Goericke, 2011; Lampe et al., 2024). Similar observations as well as the identification of dominant diatoms have been reported with microscopy (Venrick, 2002, 2009; Venrick, 2012); however, the number of diatoms that can be distinguished with light microscopy is orders of magnitude lower than estimates of their species-level diversity, and even lower when considering cryptic species (Guiry, 2024; Kooistra et al., 2007; Mann and Droop, 1996).

Compared to microscopy, sequencing of marker genes such as the 18S rRNA gene often provides higher resolution of the diatom community (Zimmermann et al., 2015). On a global scale, diatom 18S rRNA-based diversity was investigated via the *Tara Oceans* expedition, which includes few samples from coastal upwelling ecosystems, but estimated global 18S-based pelagic diatom diversity to be over 4,500 operational taxonomic units (OTUs) (Malviya et al., 2016; Pierella Karlusich et al., 2024). Although both *Tara Oceans* and separate observations from the North Atlantic Ocean show that diatom diversity remains comparatively high between the coastal and open ocean (Malviya et al., 2016; Setta et al., 2023), diatom diversity in the CCE instead exhibits a different pattern that is unique among phytoplankton groups in the region. In the CCE, diatom diversity is highest concomitantly with their abundances in the nearshore where upwelling occurs, whereas in the offshore oligotrophic region, their diversity and abundances are low (James et al., 2022; Lampe et al., 2024). These differences in diversity have also been investigated in the context of onshore and offshore eddies with different nutrient profiles (Abdala et al., 2022); however, the processes underlying these trends as well as those that govern community assembly in diatoms in upwelling regions remain underexplored. In particular, knowledge of diatom community structure may inform our understanding of carbon export as certain diatoms may disproportionally contribute to export flux, but detailed molecular-level observations of sinking diatoms in the CCE are relatively sparse (Durkin et al., 2022; Leblanc et al., 2018; Preston et al., 2020; Valencia et al., 2022).

The examination of diatom gene expression via metatranscriptomics has provided further targeted insight into diatoms' physiological responses to upwelling or other conditions found in upwelling environments. For example, diatoms have been found to proactively invest in nitrogen assimilation prior to upwelling events in order to outcompete other phytoplankton taxa (Lampe

et al., 2018b; Lin et al., 2024). Within the California Current, diatoms have also been shown to be highly transcriptionally responsive to nutrient availability and the diel cycle (Cohen et al., 2017; Kolody et al., 2019; Maniscalco et al., 2023), while employing diverse mechanisms to counteract stressors such as ocean acidification and iron stress (Kolody et al., 2022; Lampe et al., 2018a; Lampe et al., 2023). However, differences in diatom gene expression along environmental gradients that coincide with differences in their abundances and diversity in upwelling regimes remain unexplored.

Thus, here we examine patterns in pelagic diatom abundances, diversity, and gene expression in the CCE and contrast these patterns with other dominant phytoplankton groups. For each season over seven years, DNA and RNA samples for metabarcoding and metatranscriptomics respectively were collected and analyzed from California Cooperative Oceanic Fisheries Investigations (CalCOFI) surveys, hereafter referred to as the NOAA-CalCOFI Ocean Genomics (NCOG) dataset. Although the CalCOFI sampling area is largely restricted to the southern California Current, it crosses a major biogeographic boundary at Point Conception, California and contains a mixture of phytoplankton communities associated with offshore oligotrophic waters and the subarctic Northeast Pacific, enabling significant sampling of the communities associated with the overall CCE (Checkley and Barth, 2009; Venrick, 1998). Variability in phytoplankton abundances in the CCE is also most strongly in the cross-shore direction, and the sampling grid captures this variability spanning nearshore coastal samples to more offshore oligotrophic waters (Deutsch et al., 2021; Hayward and Venrick, 1998; James et al., 2022). Additionally, the four cruises per year over seven years captures seasonal and interannual variability. Overall, these data provide the basis to investigate patterns associated with the molecular-level abundances and diversity of diatoms in the CCE with relatively high

temporal and spatial resolution across wide-ranging physicochemical environments in the California Current.

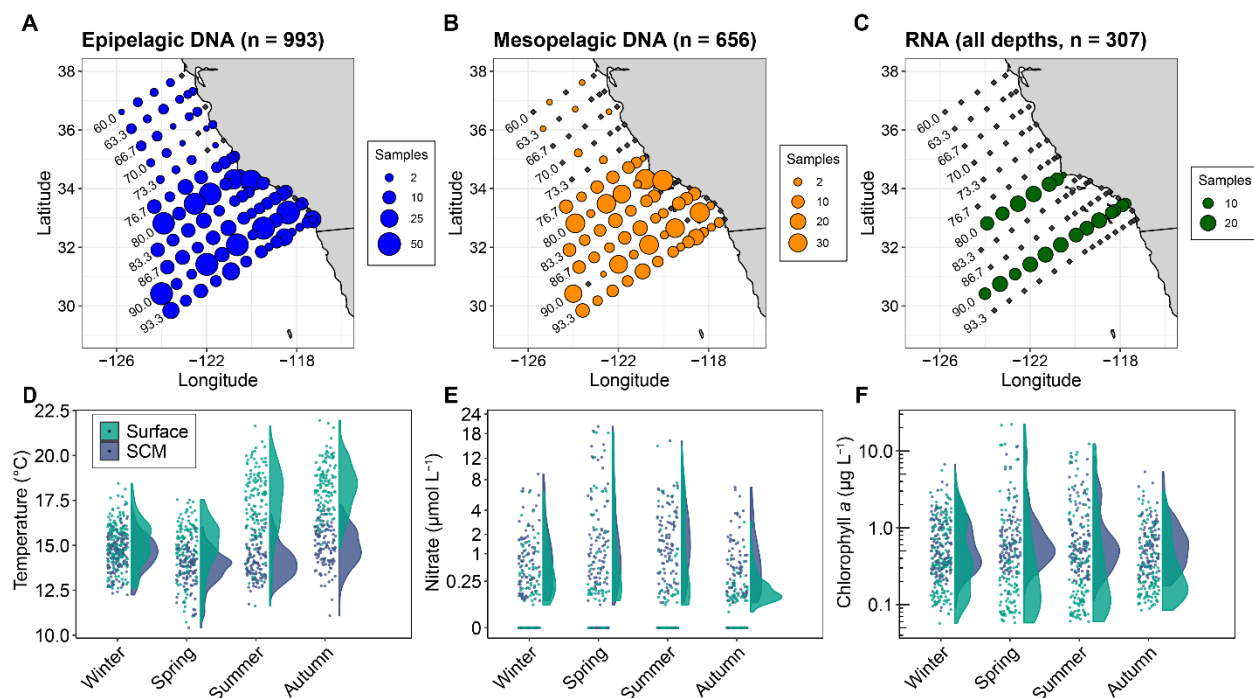


Figure 4.1: Overview of NOAA-CalCOFI Ocean Genomics (NCOG) sampling from 2014 to 2020. (A-C) Map of the CalCOFI sampling grid for epipelagic DNA, mesopelagic DNA, and RNA (all depths) respectively. Line labels correspond to the CalCOFI line and station numbering system. Bubble sizes correspond to the total number of samples collected at each station as indicated in each panel's legend. The smallest points shown as grey diamonds on lines 60.0, 66.7, 70.0, and 73.3 denote stations where no samples were taken. (D) Temperature (°C), (E) nitrate concentrations (NO₃, μmol L⁻¹), and (F) chlorophyll *a* concentrations (μg L⁻¹) across seasons corresponding to all of the 2014-2020 NCOG DNA samples. Nitrate concentrations are plotted on a cube root scale, and chlorophyll *a* concentrations are plotted on a log scale. Surface and subsurface chlorophyll max (SCM) samples are separated and coloured in green and blue respectively.

Results

Diatom abundances in the California Current Ecosystem

The updated NCOG data used here includes 993 euphotic zone DNA samples and 656 mesopelagic zone DNA samples (Figs. 4.1a,b). RNA samples for poly(A)-selected metatranscriptomics are also included with 220 euphotic zone and 87 from the mesopelagic zone

(Fig. 4.1c). From seven years of seasonal sampling and the strong cross-shore gradients in the region, the naturally high physical, chemical, and biological variability of the region was captured as observed with wide ranges in temperature, nitrate concentrations, and chlorophyll *a* concentrations (Figs. 4.1d-f) (Deutsch et al., 2021). In particular, seasonal wind-driven upwelling, primarily during spring, leads to lower temperatures coupled to higher nitrate and chlorophyll *a* concentrations (Jacox et al., 2018a). Relatively strong interannual variability was also captured due to an unprecedented marine heatwave from mid-2014 to 2016 associated with El Niño in 2015, although temperature anomalies were restricted to approximately the upper 50 m (Di Lorenzo and Mantua, 2016; Jacox et al., 2018b; Leising et al., 2015).

Metabarcoding of both the V4 and V9 hypervariable regions of the 18S rRNA gene was performed on all DNA samples to examine unique sequences, or amplicon sequence variants (ASVs). Diatoms (Bacillariophyta) in the region had 1189 and 891 ASVs for the V4 and V9 regions respectively (Fig. 4.2a). Rarefaction further suggests that the high sequencing depth achieved here enabled saturation of the total number of both V4 and V9 diatom ASVs in the region. Inference of ASVs also enables comparisons with other data sets where the same V9 primers were used (Callahan et al., 2017). Of the V9 ASVs in NCOG, 420 (47%) were undetected in other global surveys combined which includes Tara Oceans, Plankton Planet. and four stations from the Malaspina Expedition (Fig. 4.2b).

When integrated, these data offer an updated estimate of 18S-V9 diversity for pelagic diatoms in the global ocean. Assuming a log-normal (Preston) distribution of diatom ASVs and their binned abundances, a previous extrapolation from Tara Oceans suggested that there are 4,748 operational taxonomic units (OTUs) having observed 81.6% of OTUs (Malviya et al. 2016). With these combined datasets and ASV-level resolution, we now estimate that there are

11,432 diatom V9 ASVs with 62% observed (Fig. 4.2c). Importantly, ASVs represent biological DNA sequences as opposed to OTUs that are clusters of sequencing reads, offering an updated diversity estimate with intrinsic biological meaning and sequences that can be further compared in future studies (Callahan et al. 2017).

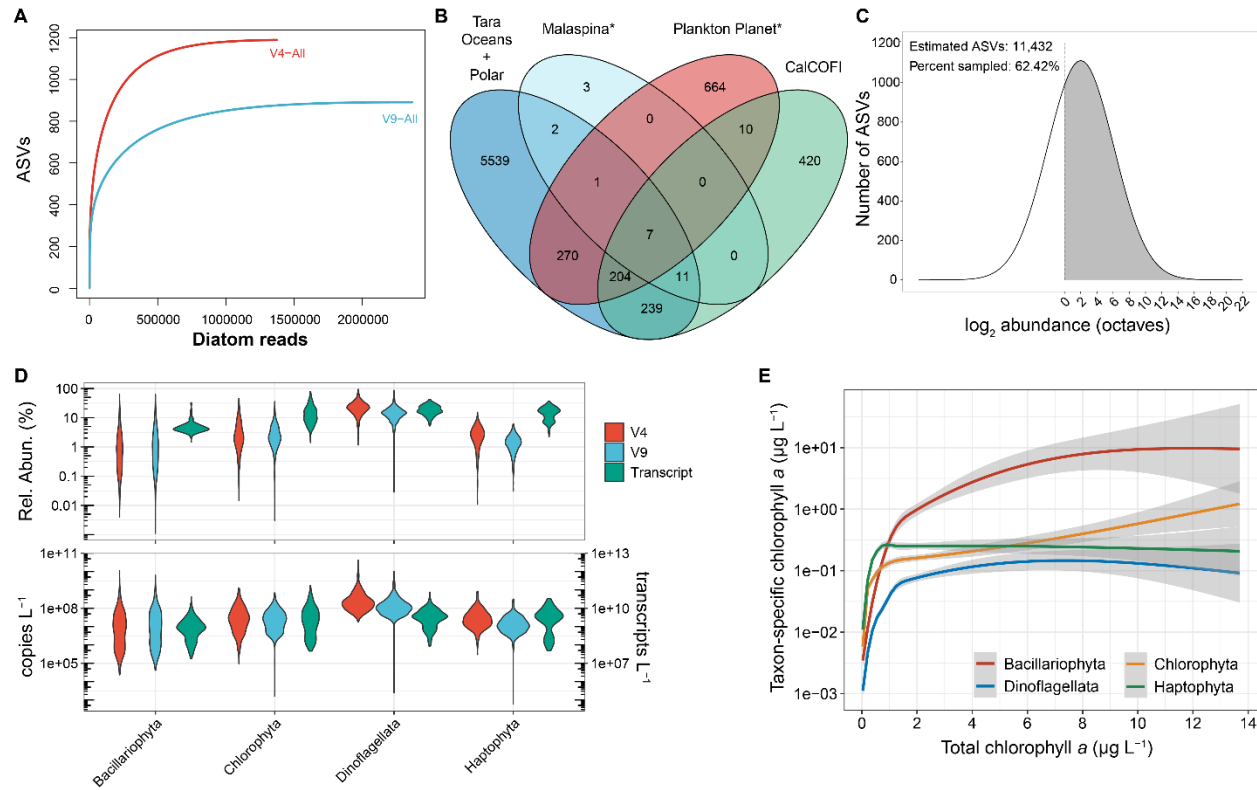


Figure 4.2: Overview of total diatom richness and abundances in the NCOG dataset from 2014 to 2020. (A) Rarefaction curves for diatom ASVs in the V4 (red) and V9 (blue) data. (B) Venn diagram of diatom 18S-V9 ASVs found within *Tara Oceans* including the *Tara Polar Circle* samples, available Malaspina Expedition and Plankton Planet samples, and CalCOFI (NCOG). (C) Log-normal distribution (Preston) of V9-18S ASV abundances from integrating different datasets shown in B. The area under the curve provides an extrapolated estimate of total number of ASVs, and therefore, an indication of the percent of ASVs sampled across datasets. (D) Relative abundances (% 18S reads or eukaryotic TPM) and absolute abundances (copies L⁻¹ or transcripts L⁻¹) of different phytoplankton groups using V4 (red), V9 (blue), or transcript abundances (green). (E) Taxon-specific biomass as chlorophyll *a* concentrations estimated with phytoclass against total chlorophyll *a* concentrations. Curves were fit with a generalized additive model.

Within the euphotic zone samples from NCOG, diatom ASVs comprised 0% to 64% of all eukaryotic taxa including non-photosynthetic organisms (Fig. 4.2d). By using an internal DNA standard, relative abundances were translated into estimates of absolute abundances for diatom ASVs for samples from all years except 2017 (*Materials and Methods*). Maxima of 1.14×10^{10} copies L⁻¹ in the V4 data (median = 6.61×10^6 copies L⁻¹) and 6.32×10^9 copies L⁻¹ in the V9 data (median = 7.47×10^6 copies L⁻¹) were observed (Fig. 4.2d). Concurrently, diatom abundances as a proportion of eukaryotic mRNA were assessed with relative and quantitative metatranscriptomics that also used internal standards in the euphotic zone samples. Relative abundances show that diatoms comprised 0.74% to 14.5% (median = 1.55%) of eukaryotic mRNA, and when converted to absolute abundances, diatoms accounted for 1.63×10^8 to 4.24×10^{10} transcripts L⁻¹ (median = 1.11×10^9 transcripts L⁻¹) (Fig. 4.2d and Supplemental Fig. 4.1).

Overall, diatom abundances between the V4 and V9 data were strongly correlated (Spearman's $\rho > 0.96$, $P < 0.0001$) with diatoms often displaying higher abundances with the V9 data (Supplemental Fig. 4.2). This bias towards diatoms with the V9 data may be due to the primers missing some eukaryotic taxa that are better captured with V4 primers (McNichol et al. 2021). Significant correlations were also observed between 18S and transcript abundances with absolute quantities showing a stronger correlation (Spearman's $\rho = 0.65$, $n = 90$) than the relative data ($\rho = 0.44$, $n = 122$) (Supplemental Fig. 4.2).

On average, diatom 18S and transcript relative abundances were lower than those of other major eukaryotic phytoplankton groups: chlorophytes (Chlorophyta), dinoflagellates (Dinoflagellata), and haptophytes (Haptophyta); however, when converted to absolute abundances, average diatom abundances were higher than chlorophytes and haptophytes (Fig. 4.2d). High dinoflagellate abundances may be due to bias from their high 18S gene copy

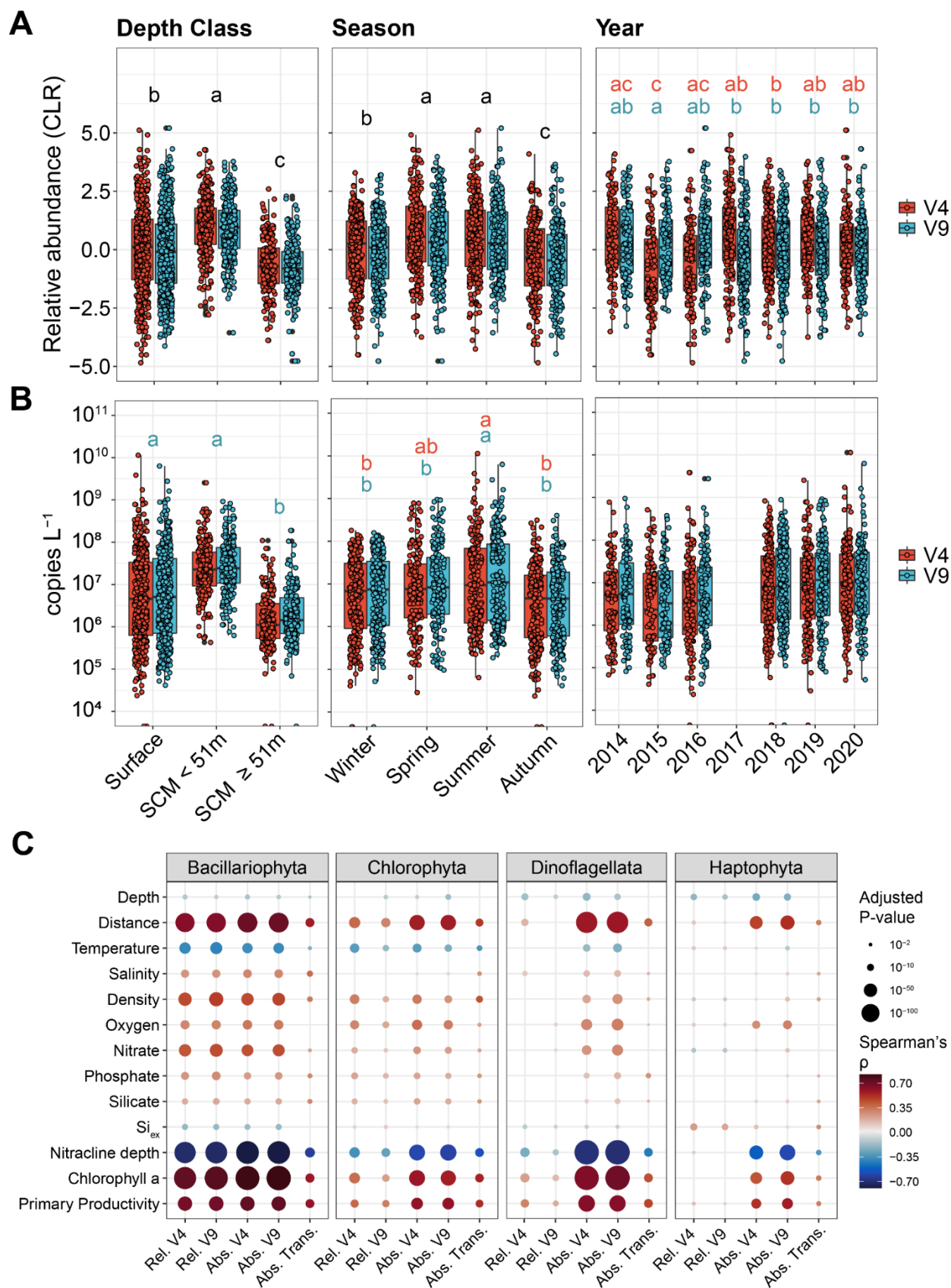
numbers (Martin et al. 2022), and their transcript abundances were more similar to those of other phytoplankton groups. Maximum values of diatom 18S abundances exceeded those of chlorophytes and haptophytes with diatom transcripts exceeding all groups (Fig. 4.2d). Similarly, diatom-specific biomass derived from HPLC pigments largely exceeded other groups besides at the lowest levels of total chlorophyll *a* (Fig. 4.2e). Overall, these wide ranges in diatom abundances and high abundances when overall phytoplankton biomass is high align with their known “bloom and bust” lifestyle (Gross 2012).

To further examine and contrast patterns in diatom abundances, relative and absolute abundances were considered against various physical, chemical, and biological measurements (Fig. 4.3). With the relative abundances, the centered log-ratio (CLR) transformation was applied to account for compositionality (Supplemental Fig. 4.3) (Gloor et al. 2017). Diatom abundances were higher in shallower subsurface chlorophyll maxima layers (SCMLs) relative to the near surface samples (normally 10 m) or SCLMs deeper than 51 m (Figs. 4.3a,b). The separation of SCMLs at 51 m was used to distinguish shallower from deeper subsurface chlorophyll maxima, the former which may exist during upwelling conditions and the latter existing under more stratified nutrient-limited conditions. In most cases, diatom abundances were also significantly higher during the spring and summer seasons compared to winter and autumn (Figs. 4.3a and 4.3b). When considering interannual variability, median diatom abundances were also lower during the 2014-2016 warm anomaly; however, these differences were not significant. As the warm anomaly was restricted to 50 m or shallower, samples solely collected within those depths were also examined, but no significant differences were observed either (Leising et al. 2015). Aggregated regional diatom cell counts from the mixed layer also do not indicate a significant difference among years (Supplemental Fig. 4.4). By correlating abundances with other measured

variables, diatom abundances were most strongly related to increased chlorophyll *a* concentrations and euphotic zone-integrated rates of primary productivity ($\text{mg C m}^{-2} \text{ d}^{-1}$) as well as lower nitracline depths and distances from shore ($\rho = 0.57\text{--}0.82$, Fig. 4.3c and Supplemental Figs. 4.5, 4.6).

These measurements with strong connections to increased diatom abundances are indicative of seasonal coastal upwelling, and align with well-known increases in diatom abundances during upwelling events. Furthermore, this pattern is distinct among eukaryotic phytoplankton groups. With 18S relative abundances and absolute transcript abundances, chlorophytes exhibited similar patterns to diatoms, albeit with weaker correlations (Fig. 4.3c). Meanwhile, dinoflagellates and haptophytes displayed weak or non-significant correlations. Higher correlations across all groups were observed when using the absolute data, which may be a result of the absolute data not being influenced by compositionality (Lampe et al. 2024a) however, diatom correlations remained the strongest. Although significant differences among seasons were also observed for other phytoplankton group abundances, there were not strong differences associated with increases in the spring and summer as observed in diatoms (Figs. 4.3 and Supplemental Fig. 4.7). This separation of diatoms compared to other groups is further evidenced by significantly lower realized temperature niches ($P < 0.0001$), further suggesting that their niche is associated with upwelling conditions (Supplemental Fig. 4.8).

Figure 4.3: Diatom abundances data against environmental measurements. (A) CLR-transformed relative abundances separated by depth class (surface, chlorophyll max < 51 m, chlorophyll $\geq$ 51 m), season, or year for the V4 (red) and V9 (blue) data. Letters denote medians that are significantly different (adjusted P -value < 0.05, Kruskal-Wallis test). If letters are black, the significant differences are the same for V4 and V9 data; otherwise, they are shown in red for V4 and blue for V9. Letters are not shown where differences are not significant. (B) Absolute abundance data (copies L⁻¹) corresponding to panel A. (C) Spearman correlation results between the abundance of dominant phytoplankton groups and continuous environmental variables. Both the relative (rel.) and absolute (abs.) abundance data are shown for both the V4 and V9 data as well absolute abundances from the metatranscriptome (transcripts L⁻¹). The color of each point corresponds to the rank correlation coefficient (ρ) whereas the size of each point corresponds to the adjusted P -value as shown in the legend. For the CLR-transformed data, correlations were computed with ALDEx2 using 1000 Monte Carlo instances. In all data sets shown, P -values were adjusted with the Benjamini-Hochberg procedure.



Diatom genera abundances and richness

To examine the composition of the diatom community, genus-level abundances and richness within euphotic zone samples were assessed (Fig. 4.4). To normalize for uneven sequencing depth across samples, samples were rarefied to 24,000 reads resulting in retainment of 938 samples (94%) with the V4 data and 988 samples (99%) with the V9 data. The V4 data had taxonomic assignments to 49 different diatom genera whereas the V9 data had 35 genera (Fig. 4.4a). Likewise, 83% of V4 reads had genus-level assignments whereas only 65% of V9 reads had genus-level assignments. Of the remaining 17% of V4 reads without genus-level assignments, 15.7% of the total community only had class level assignments, predominantly classified as polar centric diatoms (13.1%), and of the remaining 35% of V9 reads without genus-level assignments, 17.0% of the total community had class-level assignments, predominantly raphid pennate diatoms (9.4%). Eighteen genera in the V4 data with representatives from all four diatom classes were absent in the V9 data while only four genera in the V9 data absent in V4 data. Of these, *Amphora* and *Biddulphia* were detected with only one read post-rarefaction in the V4 data; similarly, one read was assigned to *Cymbella* in the V9 data which was not found in V4 data. Higher resolution taxonomic assignment with the V4 region, which is also longer than the V9 region, is likely a result of greater genetic divergence in diatoms compared to V9, particularly between araphid and raphid pennates (Malviya et al. 2016).

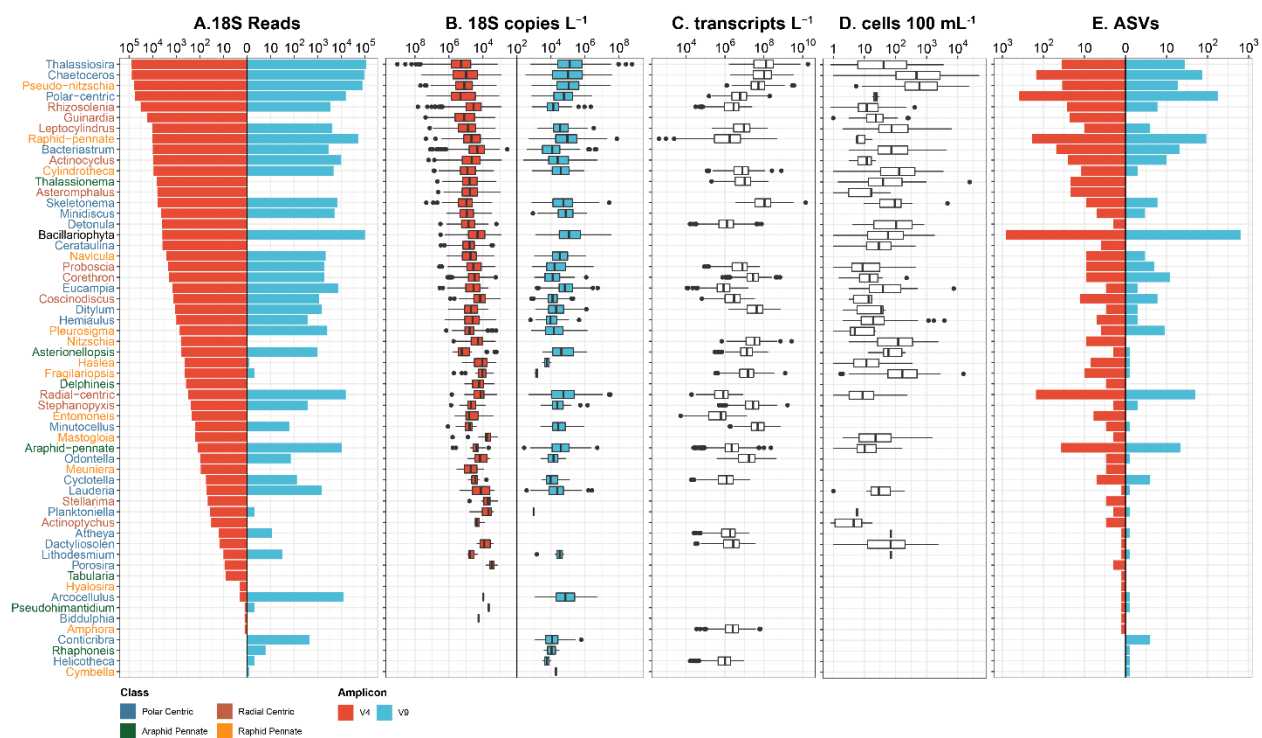


Figure 4.4: Genus-level diatom abundances and richness on log scales. Genera are color-coded by class as shown in the legend and ordered by rarified V4 read abundances. Where applicable, V4 data are shown in red and on the left, and V9 data are shown in blue on the right of each panel. ASVs without genus-level annotation are shown at the lowest taxonomic level possible. (A) Total rarified 18S read abundances (B) 18S absolute abundances (copies L^{-1}) (C) Absolute transcript abundances (transcripts L^{-1}). Genus-level assignments in the metatranscriptome that are not in the 18S data are classified on the class level. Complete genus-level transcript abundances are shown in Supplemental Figure 4.9. (D) Cell abundances from microscopy (cells $100 mL^{-1}$). Genus-level assignments in the cell counts that are not in the 18S data are classified on the class level. Both unclassified centric and pennate diatoms in the cell counts are shown as Bacillariophyta. Complete genus-level transcript abundances are shown in Supplemental Figure 4.10. (E) Genus-level richness (total number of ASVs).

However, both markers indicate that *Thalassiosira*, *Chaetoceros*, and *Pseudo-nitzschia* are the most prevalent genera in the region (Fig. 4.4). In the V4 data, these genera were 18.9%, 18.0%, and 14.2% respectively. The remaining top ten genera were *Rhizosolenia* (7.6%), *Guinardia* (4.1%), *Leptocylindrus* (2.5%), *Bacteriastrum* (2.3%), *Actinocyclus* (2.2%), *Cylindrotheca* (2.2%), and *Thalassionema* (1.5%). In the V9 data, *Thalassiosira*, *Chaetoceros*, and *Pseudo-nitzschia* comprised 20.5%, 16.8%, and 14.2% of reads respectively. As in the V4

data, *Actinocyclus* (1.8%) *Cylindrotheca* (0.9%), and *Leptocylinndrus* (0.8%) were in the top ten genera, but *Guinardia* and *Thalassionema* were completely absent in the V9 data. Rather, *Eucampia* (1.3%), *Skeletonema* (1.2%), and *Minidiscus* (0.9%) were in in the top ten V9 genera and similarly were between 0.4% and 1.5% in the V4 data. *Arcocellulus* was also one of the most abundant V9 genera (2.3%) while being one of the least abundant V4 genera.

The *Tara Oceans* expedition also found *Chaetoceros*, *Thalassiosira*, and *Pseudo-nitzschia* to be the first, second, and fourth most abundant genera with both the V4 and V9 regions (Pierella Karlusich et al. 2024). Similarly, *Actinocyclus* and *Leptocylinndrus* were also highly abundant with both amplicons and *Guinardia* was abundant in the V4 data. In contrast, *Porosira* and *Proboscia* were among the top 10 genera globally, while showing lower relative abundances in NCOG. Differences in certain genera such as the high abundances of *Fragilaria* in *Tara Oceans* while absent here are likely a result of their more limited biogeographic distribution (Pierella Karlusich et al. 2024).

Using both 18S and transcript absolute abundances, patterns remained largely the same (Figs. 4.4b,c and Supplemental Fig. 4.9). *Thalassiosira* still dominated with maximal abundances that were an order of magnitude higher than other genera. *Chaetoceros* and *Pseudo-nitzschia* also followed as highly abundant; however, *Skeletonema* exhibited some of the highest transcript abundances and based on average V4-based absolute abundances, it was within the top 10 most abundant genera, as observed with both the relative and absolute V9 data. However, differences may be from the exclusion of samples from 2017 in the absolute abundance data.

Microscopic cell counts from the near-surface also indicate that *Chaetoceros*, excluding those identified as resting spores, and *Pseudo-nitzschia* were the most abundant diatom genera (Fig. 4.4d). Similarly, dominant genera included *Thalassionema*, *Cylindrotheca*, *Leptocylinndrus*,

and *Bacteriastrum* in order of total abundance; however, *Fragilariopsis*, *Nitzschia*, *Dactyliosolen*, and *Hemialus* were the third, fourth, ninth, and tenth most abundant genera respectively while the amplicon data shows that these are rarer genera. The high abundance of *Thalassiosira* in the amplicon data contrasted with comparatively lower cell counts (twelfth most abundant) suggests that *Thalassiosira* is overrepresented with 18S sequencing, in agreement with other comparisons (Rynearson et al. 2020). Although rDNA copy number appears to be correlated with cell size (Biard et al. 2017; Zhu et al. 2005) and genome size (Martin et al. 2021), there is likely large variation in *Thalassiosira* considering their large variability in cell size (3 – 186 μm diameter) (Hasle and Syvertsen 1997) and the order of magnitude difference in copies between two *Thalassiosira* species (Zhu et al. 2005). Further discrepancies between the microscopy and amplicon data may be attributed to lack of representation in respective references, mismatched sampling in terms of location or depth for low abundance taxa such as those predominately found in SCMLs (Supplemental Fig. 4.10), or difficulty identifying small cells.

Additionally, genus-level abundances and intragenus richness (number of ASVs) were strongly correlated (Figs. 4.4e and Supplemental Fig. 4.11, $r = 0.69\text{-}0.83$). As found in *Tara Oceans*, *Chaetoceros* was not only the most abundant genus but also the most diverse (Malviya et al. 2016). The related genus *Bacteriastrum* also displayed high richness for its abundance. Other representatives of each diatom class also show deviations from the linear relationship (Supplemental Fig. 4.11). Some very low abundance genera were restricted to a single ASV which may simply be a result of low levels of sampling for those genera, although examining unrarefied data for these genera rarely indicated the presence of an additional ASV. In both datasets, the radial centric genera *Actinocyclus*, *Coscinodiscus*, and *Cyclotella* had relatively high

richness whereas the other centric genera, *Leptocylinthus* and *Eucampia* showed relative less or their abundance. However, these trends were not entirely consistent across both datasets. For example, *Pleurosigma* had high richness for its abundance in the V4 data but low richness in the V9 data (Supplemental Fig. 4.11).

Environmental and biogeochemical factors associated with diatom diversity

Alpha diversity expressed as the exponentiated Shannon's diversity index was examined within the rarefied data from the euphotic zone among both marker genes. Both the V4 and V9 data were strongly correlated, although V9 showed overall higher diversity and richness but lower evenness (Supplemental Fig. 4.12) (Hill 1973). To examine influences on alpha diversity, differences in Shannon diversity values were compared with the other physical, chemical, and biological measurements (Fig. 4.5 and Supplemental Fig. 4.13). As observed with the diatom abundances, increases in diversity aligned with signatures of upwelling: the spring season, high chlorophyll *a* concentrations, high primary productivity, low distances from shore, and shallow nitracline depths (Kruskal-Wallis test, Figs. 4.5a,b). Diversity was also higher in shallower SCMLs (Fig. 4.5a). Nitracline depth, distance from shore, and primary productivity similarly explained the greatest amount of deviance with generalized additive models (13.0-15.7%), with the explained deviance by chlorophyll *a* concentrations being lower (4.3-5.4%) (Supplemental Fig. 4.13). However, a previous analysis using a subset of these data and corresponding phytoplankton pigment concentrations indicates that diatom-specific chlorophyll *a* accounts for a greater amount of deviance (15-20%) (Lampe et al. 2024a). Moreover, diatom diversity displays a right-shifted unimodal trend and richness increases with abundance, indicating that diatom community evenness declines at high diatom abundances (Supplemental Fig. 4.14). All other

examined variables here displayed weak relationships, and interannual variability differed between the two amplicons as observed with the relative abundances (Fig. 4.5 and Supplemental Fig. 4.13).

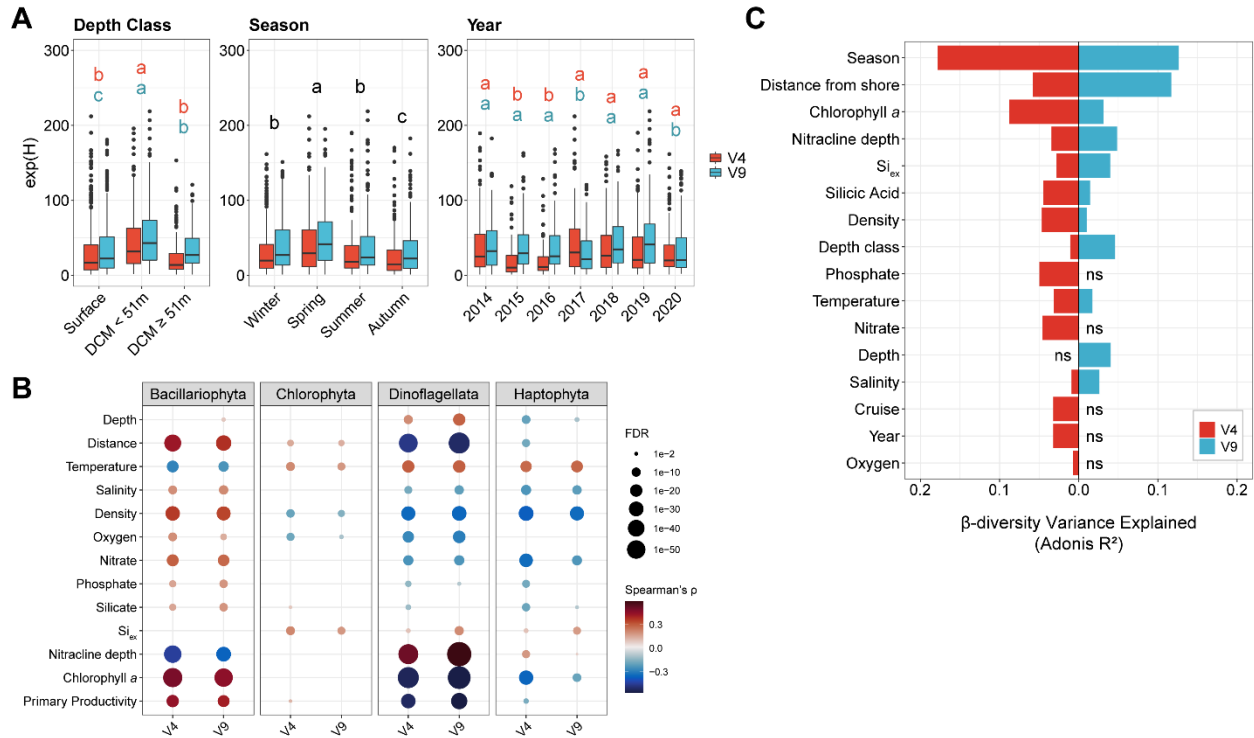


Figure 4.5: Exponentiated Shannon diversity [exp(H)] compared to environmental measurements. (A) Shannon diversity separated by depth class (surface, chlorophyll max < 51 m, chlorophyll $\geq$ 51 m), season, or year. Letters denote medians that are significantly different ($P < 0.05$, Kruskal-Wallis test). If black, the significant differences are the same for V4 and V9 data, otherwise, they are shown in red for V4 and blue for V9. (B) Spearman correlation results between Shannon diversity within dominant phytoplankton groups and continuous environmental variables for both the V4 and V9 data. (C) The percentage of variance in diatom beta diversity explained by different environmental variables.

These trends further contrast diatoms with other phytoplankton groups (Fig. 4.5b). In chlorophytes, correlations were largely not significant, and in haptophytes the strengths of the correlations were weaker and negative. Meanwhile, dinoflagellates displayed a strongly opposing pattern where increased diversity aligned with conditions opposite of those for diatoms, further supporting their opposing niches in the CCE (Lampe et al. 2024a).

The influence of environmental conditions on differences within the diatom community, i.e. beta diversity, were further assessed with PERMANOVAs (adonis) on robust Aitchison distances, another compositional data analysis approach based on CLR-transformed counts for beta diversity (Martino et al. 2019). After filtering for only samples with all environmental variables and abundance thresholds (*Materials and Methods*), 416 samples in the V4 data and 535 samples in the V9 data were used. Seasonality was the strongest individual explanatory variable of diatom community dissimilarity for both marker genes ($R^2 = 17.8\%$ [V4] and 12.6% [V9]) in agreement with the patterns of seasonal upwelling aligning with changes in abundances and alpha diversity (Fig. 4.5c). Other variables, however, were less explanatory. Distance from shore closely followed seasons in the V9 data ($R^2 = 11.7\%$) but to a lesser degree in the V4 data ($R^2 = 5.8\%$). All other variables, including nitracline depth and chlorophyll *a* individually accounted for less than 10% of the explained variance with both amplicons. Distance between stations also did not strongly correlate with community dissimilarity when considering either all samples (Mantel test, $r = 0.05$ [V4] and 0.06 [V9], $P < 0.05$) or individual cruises as only three out of the 27 cruises had significant correlations (Mantel test, $r = 0.28$ - 0.29 , $FDR = 0.027$, Supplemental Fig. 4.15).

Biogeography of individual diatom ASVs

The absence of a strong correlation between sample and community distances could be explained by diatoms' high capacity for dispersal and the relatively small size of the study area (Cermeno and Falkowski 2009). However, it has also been suggested that diatoms can be dispersal limited (Vanormelingen et al. 2007). Considering the general increase in diatom diversity nearshore (Fig. 4.5 and Supplemental Fig. 4.13) as well as the frequently observed

steep cross-shore physicochemical and biological gradients (Deutsch et al. 2021, James et al. 2022; Matthews and Ohman 2023), the biogeography of individual diatom ASVs was investigated with abundance-occupancy distributions. Here, the average relative abundance was used with occupancy defined as the percentage of stations where each ASV was detected at least once (Fig. 4.6 and Supplemental Fig. 4.11). Assuming unlimited dispersal, neutral models predict a relationship between abundance and occupancy where taxa that are highly abundant are also expected to be widespread (Shade and Stopnisek 2019; Sloan et al. 2006). These models can be fit to abundance-occupancy distributions, and taxa that deviate below the model predictions (95% confidence interval) exhibit low occurrence for their abundance, implying that they may be dispersal limited (Shade and Stopnisek 2019).

Overall, the abundance and occurrence of eukaryotic ASVs were well described by the neutral model ($\rho = 0.71-0.73$; Supplemental Fig. 4.16). Within diatoms, station occupancy was as high as 63 out of 66 stations in the V4 data and all 66 stations in the V9 data, but more than 88% of all diatom ASVs were only found in fewer than half of the stations. Other phytoplankton groups had ASVs detected in all stations with both amplicons, but most of their ASVs were also only detected in fewer than 50% of the stations.

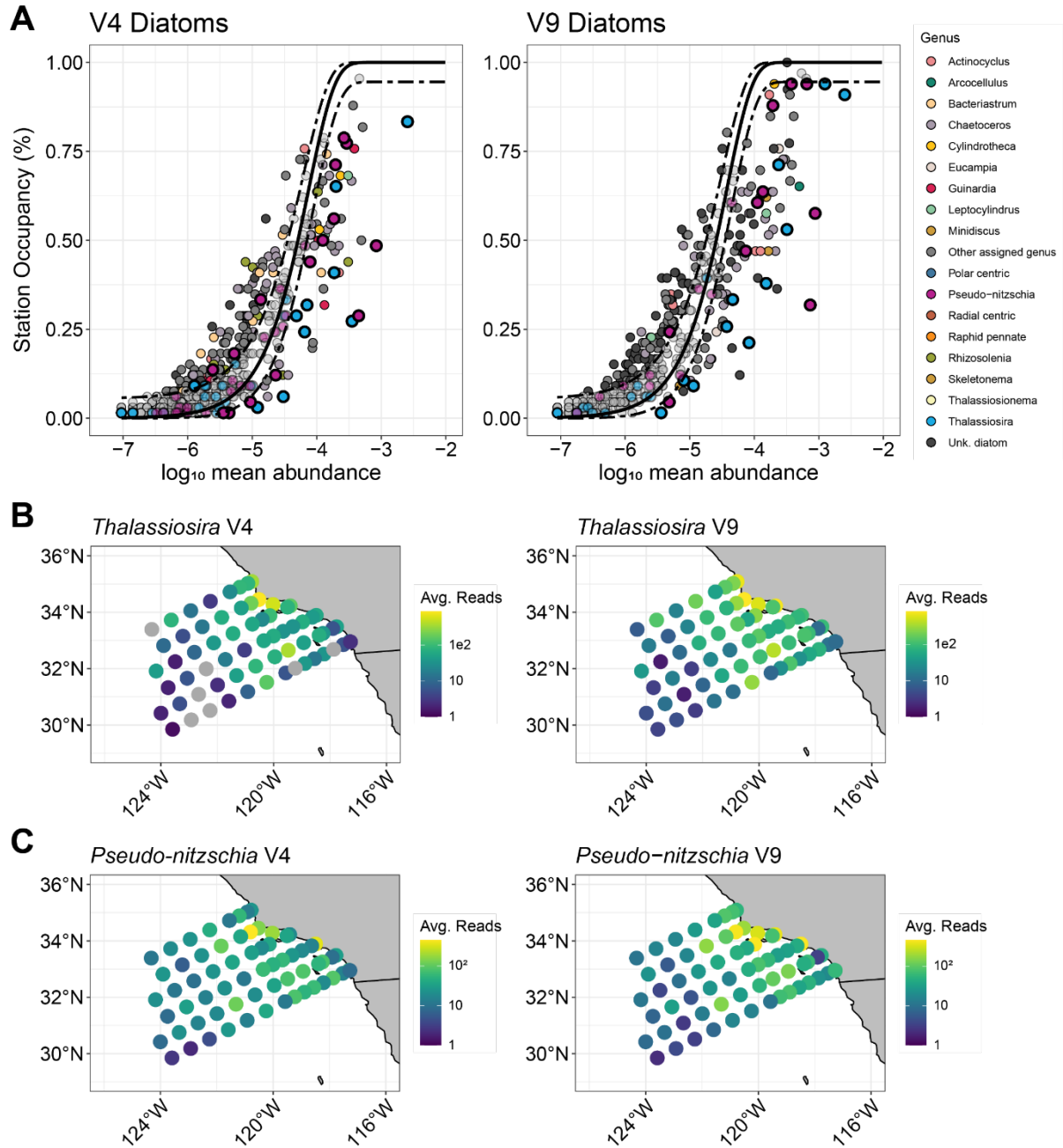


Figure 4.6: (A) Abundance-occupancy distributions fit with a neutral model for the V4 (left) and V9 (right) datasets. Abundances are the average rarefied abundance in the dataset while the occupancy is based on the percentage of stations where the ASV was detected. ASVs that occur more or less frequently than predicted by the model are shown above or below the 95% confidence intervals (dashed lines) around the model prediction (solid black line). These ASVs that deviate from the neutral model are colored by genus or the lowest possible taxonomic classification where genus-level classification is not available. The full data from which the neutral model was fit is shown in Fig. S15. (B and C) The average read abundance on a log scale per station for *Thalassiosira* and *Pseudo-nitzschia* ASVs that are predicted to be dispersal limited by the neutral model.

Over 8% and 15% of diatom ASVs in the V4 and V9 data respectively were found to be below the model prediction, which is higher than other phytoplankton groups (3-5%) and suggests dispersal limitation (Fig 4.6a). Moreover, *Thalassiosira* and *Pseudo-nitzschia* ASVs were significantly enriched in this dispersal limited fraction (Fisher's exact test, FDR < 0.06) and displayed a trend of higher abundances in the nearshore environment (Figs. 4.6b,c). Although *Chaetoceros* ASVs were not significantly overrepresented in the dispersal limited group, many *Chaetoceros* ASVs also exhibited this same trend (Supplemental Fig. 4.17). Overall, these ASVs are not only among the most abundant individually but also from the most prominent genera. This pattern, however, indicates that they are geographically limited within the study region and frequently either undetected or found in low abundances in the offshore region. As successful colonization offshore requires both arrival and establishment (Barberán et al. 2014), this suggests that they are either removed from the euphotic zone by various processes or that colonization in the offshore region is short-lived. Indeed, the steep nutrient gradients in the cross-shore direction leads to nitrogen limitation, iron limitation, or increased silicon stress which may result in senescence and sinking (Deutsch et al. 2021, Forsch et al. 2023). Cells may also be removed from the euphotic zone by physical processes, i.e. subduction (Stukel et al. 2023). Alternatively, they may be infected or grazed either before they can disperse or become established.

To investigate potential connections to other organisms that may cause cell death and prevent dispersal, the co-occurrence of diatom ASVs with other taxa was examined, including prokaryotic ASVs from 16S amplicon sequencing (*Materials and Methods*). Overall, diatom ASVs both above and below the neutral model exhibited similar co-occurrence profiles (Supplemental Fig. 4.18). Microbial grazers represent the main biological control on phytoplankton abundances (Kuhlich et al. 2024); however, there was no clear increase for

dispersal limited diatom ASVs in the relative number of connections with ciliates, metazoans, copepods, radiolarians, or choanoflagellates. Diatoms may also be infected by parasites belonging to the Syndiniales, Oomycota, and Chytridiomycota lineages (Kuhlisch et al. 2024), and of these parasites, only Oomycota connections were elevated for dispersal limited diatoms. However, Oomycota connections represented less than 0.5% of all diatom connections. Diverse bacteria also cause pathogenicity in diatoms including the grammaproteobacter *Croceibacter atlanticus*, although no difference was observed for connections to this species (Bartolek et al. 2022; van Tol et al. 2016). Other known pathogenic bacteria were largely undetected, which may result from the inability to resolve species with 16S metabarcoding (Branscombe et al. 2024; Kuhlisch et al. 2024). Although co-occurrence may not detect a true interaction and there are other interactions that are not accounted for such as viral infection, these results suggest that differential biological interactions do not account for differences in diatoms' ability to disperse (Röttgers and Faust 2018).

Diatoms in the mesopelagic zone

Samples from 170-200 m and 515 m were examined to further explore diatoms below the euphotic zone, particularly those that may be dispersal limited from sinking or subduction (Fig. 4.1b) (Stukel et al. 2023). In comparison, the deepest chlorophyll maximum sample used here was found at 130 m at the most offshore station (station 090.0 120.0) where the euphotic zone is normally much deeper than in the nearshore environment. Overall, the CCE exhibits above average carbon flux, and diatoms often represent a high proportion of the community within sinking particles (Durkin et al. 2022; Krause et al. 2015; Preston et al. 2020; Stukel et al. 2023; Valencia et al. 2022). The depths sampled here are below those considered to be upwelling

source waters in the region; therefore, diatoms found in the mesopelagic in the here may significantly contribute to carbon sequestration (Fiechter and Moore 2024).

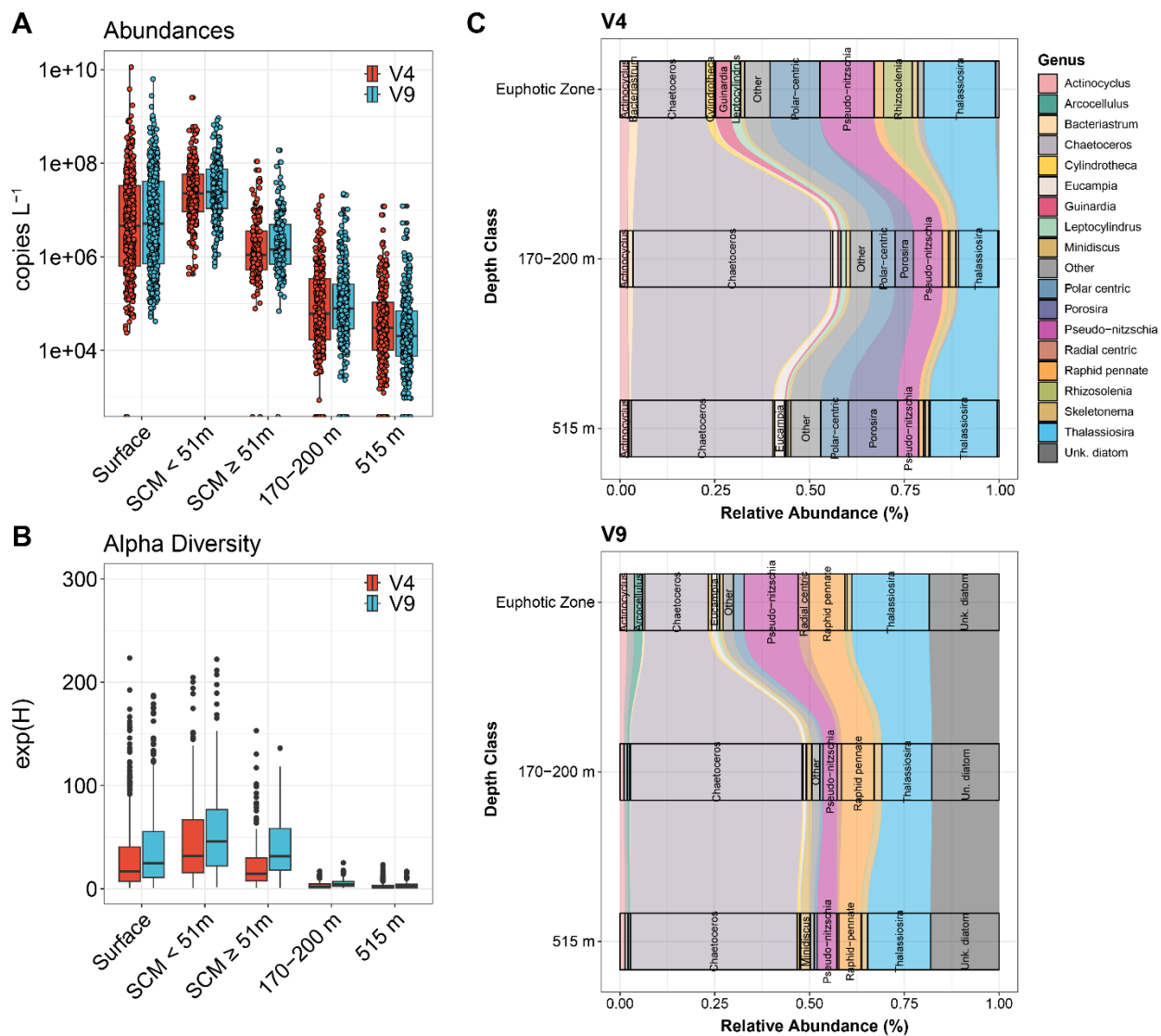


Figure 4.7: Diatom abundances and diversity in the mesopelagic zone. Absolute abundances (copies L^{-1}) by depth class when including mesopelagic depths. (B) Exponentiated Shannon diversity [exp(H)] by depth class when including mesopelagic depths. (C) Relative abundances of the most abundant diatom genera or the lowest possible taxonomic classification where genus-level classification is not available within the euphotic zone, 170–200m, and 515 m for V4 (top) and V9 (bottom).

On average, diatoms were less than 0.5% of the eukaryotic community in the mesopelagic samples but were as high as 7.6% (V4) or 17.6% (V9), which largely exceeded chlorophytes and haptophytes. Dinoflagellates, however, were on average 7-10% of the community and as high as 71-91%. When using absolute abundances, diatoms generally declined among depth categories from shallow SCMLs to the 515 m samples (Fig. 4.7a). Although diatom transcript abundances were also lower in the mesopelagic, active transcription implies that some cells were living and active, perhaps having recently sunk out of the euphotic zone (Supplemental Fig. 4.1).

Diatom richness in the mesopelagic was 515 V4 and 409 V9 ASVs, which is similarly lower as the euphotic zone samples had 882 V4 and 719 V9 ASVs. Alpha diversity expressed as the Shannon index was also significantly lower at these depths (Kruskal Wallis, $P < 0.05$) (Fig. 4.7b). Within the diatom ASVs predicted to be dispersal limited by the neutral model, 88% of V4 ASVs and 92% of V9 ASVs were detected in the mesopelagic samples, suggesting that these ASVs are sinking out of the euphotic zone before dispersing or becoming established in the offshore region.

Within the diatom ASVs in the mesopelagic zone however, 42-56% were undetected in euphotic zone. These ASVs include five genera in the V4 data and one in the V9 data that were undetected in the euphotic zone, indicating that there may be remaining undetected diversity in the euphotic zone or that ASVs from outside the study region enter the mesopelagic zone via lateral advection (Supplemental Fig. 4.19). Conversely, 14 genera in the V4 data and nine in the V9 data detected in the euphotic zone were not found in the mesopelagic samples (Supplemental Fig. 4.19). These genera were only found in low abundances in the euphotic zone, and therefore, they may not significantly contribute to export.

As in the euphotic zone, *Thalassiosira*, *Chaetoceros*, and *Pseudo-nitzschia* were the most abundance diatom genera in the mesopelagic zone, further indicating that these dominant genera sink out of the euphotic zone rather than disperse offshore (Fig. 4.7c and Supplemental Fig. 4.11). However, *Chaetoceros* rather than *Thalassiosira* dominated the samples at depth. *Thalassiosira* and *Pseudo-nitzschia* have previously been linked to sinking particle flux in the region, but these results suggest that *Chaetoceros* may also play a large role (Preston et al. 2020; Sekula-Wood et al. 2009; Valencia et al. 2022). *Skeletonema* and *Actinocyclus* also maintained their high relative abundances from the euphotic zone to the mesopelagic zone. In contrast, some genera that were less abundant in the euphotic zone were instead the most abundant genera in the mesopelagic zone, although there were differences among the two amplicons with the V9 data having a relatively large proportion of unclassified ASVs. In the V4 data, *Eucampia* was the 19th most abundant genus, but continued to increase in relative abundances with depth to become the 6th most abundant genus. Furthermore, *Porosira* was one of the least abundant in the euphotic zone, but became one of the most abundant, exceeding *Pseudo-nitzschia* on a relative basis and having the highest average absolute abundances. In the V9 data, *Minidiscus* was found to have relatively high abundances matching previous observations within sinking particles in the California Current and Mediterranean Sea (Durkin et al. 2022; Leblanc et al. 2018).

Diatom gene expression in the California Current Ecosystem

From the RNA samples, which span onshore to offshore on two distinct sampling lines (Fig. 4.1c), assembled protein coding sequences that were taxonomically annotated as diatoms were further annotated with both the KEGG database. Sequences were then clustered into orthologous groups resulting in 26,616 unique diatom gene clusters. Samples were also only collected near local apparent noon to minimize effects from the diel cycle.

Within the euphotic zone, the abundances diatom gene clusters on an absolute basis varied many orders of magnitude (Supplemental Fig. 4.20). Approximately 66% of these clusters had no associated functional annotation; however, diverse cellular processes were represented among those that had annotations (Supplemental Fig. 4.20). Overall, photosynthesis was the most highly expressed functional category with the photosystem II genes, *psbA* and *psbD*, as the most abundant gene clusters overall. Following, the RuBisCo genes (*rbcL* and *rbcS*), which catalyze the first step in the Calvin-Benson cycle for carbon fixation, were among the next most abundant. As diatoms largely have an obligate requirement for silicon for their cell walls (frustules), silicic acid transporter (*SIT*) transcripts were also highly abundant, along with transcripts for frustulin, which may help protect the frustule (Kröger and Poulsen 2008). Other prominent nutrient acquisition gene transcripts include those for nitrate transporters (*NRT*) and phytotransferrin (*pTF* or *ISIP2a*), which mediates inorganic ferric iron uptake (McQuaid et al. 2018). Xanthorhodopsin which may aid diatoms in storing energy in the vacuole as well as genes associated with the early stages of sexual reproduction (*Sig1-3*) were also highly abundant (Armbrust 1999; Sunda and Marchetti 2024).

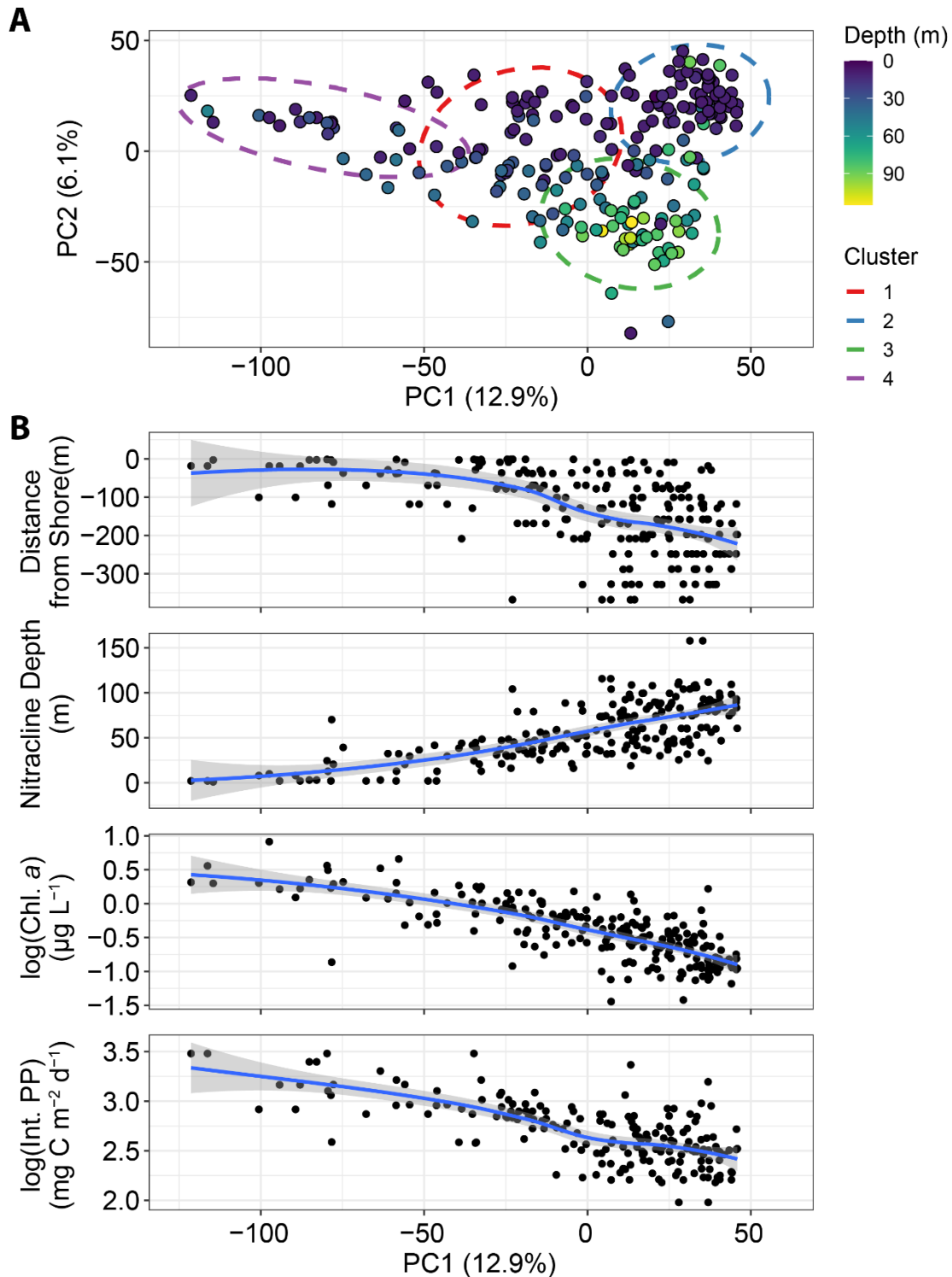


Figure 4.8: Diatom gene expression. (A) PCA of diatom expression with k-means clusters. Colors correspond to the depth of each sample. (B) Distance from shore (m), nitracline depth (m), chlorophyll *a* on ($\mu\text{g L}^{-1}$) on a log scale, and integrated primary productivity ($\text{mg C m}^{-2} \text{ d}^{-1}$) on a log scale plotted against the first principal component. Lines represent generalized additive models with associated 95% confidence intervals.

Overall, diatom transcription across all euphotic zone samples coalesced into four clusters that separated along the environmental conditions matching those associated with increased abundances and diversity (Fig. 4.8). This pattern in gene expression also differed from that of chlorophytes and haptophytes; however, dinoflagellate expression also separated along these same variables (Supplemental Fig. 4.11), aligning with dinoflagellates' opposing patterns to diatoms as observed with diversity (Fig. 5). One cluster in diatoms in particular, shown as Cluster 4, comprised samples that correspond with the same upwelling-related conditions as abundances and diversity: nearshore, shallow nitracline depths, high chlorophyll *a* concentrations, and high rates of primary productivity (Fig. 4.8). Furthermore, 92% of the samples in this cluster were from the spring and summer seasons, where these seasons only accounted for 41% to 49% of the remaining clusters' samples. In contrast, Clusters 2 and 3 correspond to conditions associated with the offshore oligotrophic region and display separation along the second principal component, which corresponds with depth: Cluster 2 represents offshore shallow samples, and Cluster 3 represents offshore deeper samples (Fig. 4.8).

Genes that are overrepresented in cluster 4 are indicative of relative increases in expression within the diatom community during coastal upwelling conditions. As these conditions also correspond to when diatom abundances are high, many of the most abundant genes are found to be overrepresented in this cluster including those associated with photosynthesis, carbon fixation, and nutrient uptake include *NRT*, *SIT*, *pTF*, and the other inorganic iron transporter, *FTR* (Lampe et al. 2024b). In contrast, the sexual reproduction genes, *SigI-3*, were more abundant in the other clusters indicating a potential shift from asexual reproduction to sexual reproduction as upwelling subsides.

Discussion

By surveying phytoplankton groups within the CCE, we observed that the regional diatom community exhibits unique patterns linked to seasonal coastal upwelling. In agreement with their well-known increased abundances when coastal upwelling arises (Estrada and Blasco 1985), diatom abundances increased in the spring and summer seasons, when upwelling-favorable winds are most likely to occur, and in the nearshore environment with shallow nitracline depths (Fig. 4.3) (Chavez and Messié 2009). Overall, diatoms dominate when phytoplankton biomass is highest, further showing strong positive correlations with chlorophyll *a* concentrations and rates of primary productivity (Figs. 4.1 and 4.3). These seasonal upwelling conditions also corresponded to relative increases in expression for genes associated with growth, such as those involved in photosynthesis, carbon fixation, and nutrient uptake (Fig. 4.8). Considering that cyanobacteria, as an early photosynthetic ancestor to diatoms, modulate their physiology based on seasonal changes even though their generation times are significantly shorter, these seasonal patterns raise the question as to whether diatoms may also seasonally alter their physiology in anticipation of coastal upwelling or other seasonal changes (Jabbur et al. 2024).

However, these increases in abundances do not lead to an overwhelming dominance of certain bloom-forming diatoms. Rather, diatom diversity increases under the same conditions that lead to increased abundances, indicating that seasonal coastal upwelling leads to a more diverse diatom community. Indeed, seasonality was found to explain the greatest amount of variance in diatom community dissimilarity (Fig. 4.5c).

This strong seasonal pattern and increased diversity with upwelling conditions was unmatched by other groups, besides dinoflagellates which exhibited contrasting diversity trends

(Fig. 4.5b and Supplemental Fig. 4.7) (James et al. 2022). As previously observed with taxon-specific biomass in the region, dinoflagellate diversity is highest when their biomass is lowest, whereas diatom diversity is highest when their biomass is highest (Lampe et al. 2024a).

Considering these divergent diversity trends and a significant difference in realized temperature niches (Supplemental Fig. 4.8), these results further reinforce the classical view of diatoms and dinoflagellates on opposite ends of the r- vs K-selected continuum, where diverse diatoms proliferate in the turbulent nutrient-rich upwelling region and dinoflagellates dominate in the offshore oligotrophic region as they are likely bolstered by mixotrophy (Margalef 1978; Stoecker et al. 2017).

Certainly, there are diatoms that exhibit unique strategies to grow in oligotrophic conditions, which may support observations of similar levels of diatom diversity between the coastal and open ocean (Kemp and Villareal 2018; Malviya et al. 2016; Setta et al. 2023). These observations will need to be reconciled with the steep decline in diversity toward the oligotrophic region observed here, particularly by including observations from other coastal upwelling systems. Globally, diatom diversity is lowest in the polar regions, but exhibits a wide range in other latitudes (Pierella Karlusich et al. 2024). Considering these observations and the high level of diversity observed here that was undetected in global surveys, it is possible that coastal upwelling regions harbor some of the highest levels of diatom diversity in the ocean (Fig. 4.2b).

The absence of these taxa in global surveys may, in part, be attributed to dispersal limitation for diatoms that are abundant in the nearshore coastal upwelling zone. Notably, many abundant diatom ASVs belonging to the genera *Thalassiosira*, *Chaetoceros*, and *Pseudo-nitzschia* appear to be restricted to the nearshore environment (Fig. 4.6 and Supplemental Fig. 4.17). Their absence in the offshore region contrasts with their high level of detection in the

mesopelagic zone, indicating that these taxa, among other dispersal limited diatoms, are blooming in the nearshore environment and sinking out of the euphotic zone rather than being advected or becoming established offshore. Along with certain other diatoms that are prevalent in the mesopelagic zone, these diatoms may disproportionality contribute to carbon export.

Although seasonality was consistently associated with differences in diatom abundances, diversity, and gene expression, patterns were inconsistent when considering interannual variability. Notably, this time series encompasses an unprecedented marine heatwave from mid-2014 to 2016 associated with El Niño in 2015, but significant differences among years were not observed (Di Lorenzo and Mantua 2016; Jacox et al. 2018; Leising et al. 2015). However, these data remain almost exclusively within the positive or “warm” phase of the Pacific Decadal Oscillation (PDO) and the negative phase of the North Pacific Gyre Oscillation (NPGO) indices, which have been correlated with regional physical and biological fluctuations (Werb and Rudnick 2023, Di Lorenzo et al. 2008). Thus, longer-term observations that span greater variability in these indices may reveal patterns in interannual or decadal variability for diatoms in the California Current.

Materials and Methods

Sample collection and environmental variables

Samples were collected from select stations on seasonal CalCOFI cruises that occurred from 2014 to 2020 (Fig. 4.1a). CalCOFI surveys the California Current ecosystem in a standardized grid pattern with most sampled stations occurring the southern California Current region. Winter cruises occurred during January and February, spring cruises occurred during April, summer cruises occurred in July and August, and autumn cruises occurred in October and

November. There was no spring cruise in 2020 due to the COVID-19 pandemic. There are also fewer samples from the northernmost lines (lines 60.0 through 73.3, Fig. 4.1a) as those were only visited on certain Winter and Spring cruises; northern line samples used here were collected on the Spring 2017 cruise (1704SH) and the Winter 2020 cruise (2001RL).

At each station, seawater was collected targeting four depths when bathymetry permitted: the near-surface (normally 10 m), the subsurface chlorophyll maximum layer (SCML), either 170 m or 200 m, and 515 m with a CTD rosette. Chlorophyll *a* was collected via gentle vacuum filtration of seawater onto GF/F filters then measured on a Turner 10-AU fluorometer via the acidification method (Strickland and Parsons 1972). Dissolved oxygen was measured using the Carpenter modification of the Winkler method (Carpenter 1965). Macronutrient concentrations were measured on a QuAatro continuous segmented flow autoanalyzer (SEAL Analytical) alongside reference materials (KANSO technos) (Armstrong et al. 1967; Gordon et al. 1992; Murphy and Riley 1962). The nitracline depth was then defined as the depth at which nitrate concentrations reach $1 \mu\text{mol L}^{-1}$. Si_{ex} , a biogeochemical proxy that suggests Fe limitation in diatoms due to preferential utilization of silicic acid (Hogle et al. 2018; Hutchins and Bruland 1998; King and Barbeau 2011), was then calculated as described by Hogle et al. (2018). Briefly, $Si_{ex} = [\mu\text{mol } H_4SiO_4 \text{ L}^{-1}] - ([\mu\text{mol } NO_3 \text{ L}^{-1}] * R_{Si:NO_3})$ where $R_{Si:NO_3}$ is the preformed ratio of silicic acid and nitrate from upwelled source waters. Based on averaged values from the upwelling source water densities in the region ($\sigma_\theta = 25.8 \text{ kg m}^{-3}$), $R_{Si:NO_3} = 0.99$ (Hogle et al. 2018).

Primary productivity was estimated via ^{14}C uptake at select stations. Briefly, seawater was collected from six depths representing 56%, 30%, 10%, 3%, 1%, and 0.3% surface light levels shortly before local apparent noon and dispensed into triplicate 250 mL polycarbonate

bottles (two light bottles and one dark control). Bottles were then inoculated with $\text{NaH}^{14}\text{CO}_3$ and incubated until civil twilight in tubes with flow-through seawater and neutral-density screens to simulate *in situ* light levels. Following incubation, samples were filtered onto 0.45 μm HA filters (Millipore), acidified with HCl, immersed in scintillation fluor, and measured with a scintillation counter once back onshore. Half light-day productivity at each depth was averaged between the two light bottles and corrected with the dark uptake bottle then multiplied by 1.8 to obtain 24 hour productivity (Eppley 1992).

Nucleic acid sample collection and extraction

For DNA samples, 0.20 to 8.85 L (mean = 4.16 L) of seawater was filtered from onto 0.22 μm Sterivex™ filters. RNA samples were collected simultaneously following the same approach but with generally higher volumes (2.00-8.67 L, mean = 4.44 L). RNA samples were also only collected near local noon to avoid bias from the diel cycle. Following filtration, samples were immediately flash frozen in liquid nitrogen, then stored at -80°C post-cruise. DNA was extracted with the Macherey-Nagel NucleoMag Plant kit on an Eppendorf epMotion 5075TMX and assessed on a 1.8% agarose gel. At the start of DNA extraction (addition of lysis buffer), 1.74 to 3.78 ng of *Schizosaccharomyces pombe* genomic DNA was added to each sample as an internal standard for all years except euphotic zone samples from 2017 (Lin et al. 2018). RNA was extracted with the Macherey-Nagel NucleoMag RNA kit on the Eppendorf epMotion. As an internal standard, 2-5 billion copies of Invitrogen™ ArrayControl™ RNA Spikes #1 and #8 were added to the lysis buffer of each sample at a 2.66:1 ratio.

Amplicon library preparation and sequencing

Amplicon libraries separately targeting the V4 and V9 regions of the 18S rRNA gene and the V4-V5 region of the 16S rRNA gene and were constructed via a one-step PCR with the TruFi DNA Polymerase PCR kit to simultaneously amplify the region of interest and barcode each sample. For 18S-V4, the V4F (5'-CCA GCA SCY GCG GTA ATT CC-3') and V4RB (5'-CCA GCA SCY GCG GTA ATT CC-3') primer set was used (Berdjeb et al. 2018). For 18S-V9, the 1389F (5'-TTG TAC ACA CCG CCC-3') and 1510R (5'-CCT TCY GCA GGT TCA CCT AC-5') primer set was used (Amaral-Zettler et al. 2009). For 16S, the 515F-Y (5'-GTG YCA GCM GCC GCG GTA A-3') and 926R (5'-CCG YCA ATT YMT TTR AGT TT-3') primer set was used (Parada et al. 2016).

Each reaction was performed with an initial denaturing step at 95°C for 1 minute followed by 30 cycles of 95°C for 15 seconds, 56°C for 15 seconds, and 72°C for 30 seconds. 2.5 µL of each PCR reaction was run on a 1.8% agarose gel to confirm amplification, then PCR products were purified with Beckman Coulter AMPure XP beads following the manufacturer's instructions. PCR quantification was performed in duplicate using the Invitrogen Quant-iT PicoGreen dsDNA Assay kit. Samples were then combined in equal proportions into multiple pools followed by another 0.8x AMPure XP bead purification on the final pool. DNA quality of each pool was evaluated on an Agilent 2200 TapeStation, and quantification was performed with the Qubit HS dsDNA kit. Each 16S or 18S pool was sequenced on an Illumina MiSeq (2 x 300 bp for 16S and V4 or 2 x 150 bp for V9) except for the one pool for the 2014-2016 euphotic zone V9 samples, which was run on an Illumina NextSeq 500 (Mid Output, 2 x 150 bp).

Analysis of amplicon sequences

Amplicons were first analyzed with QIIME2 v2019.10 (Bolyen et al. 2019). Briefly, paired-end reads were trimmed to remove adapter and primer sequences with cutadapt (Martin 2011). Trimmed reads were then denoised with DADA2 to produce amplicon sequence variants (ASVs; maxEE = 2, chimera-method = “pooled”). Each MiSeq run was denoised with DADA2 separately to account for different error profiles in each run then merged. Taxonomic annotation of ASVs was performed with the q2-feature-classifier naïve bayes classifier and the PR² database (v4.13.0) (Bokulich et al. 2018; Guillou et al. 2012; Pedregosa et al. 2011). When examining dinoflagellates (Dinoflagellata), ASVs assigned to the Syndiniales class were excluded.

To compare the presence or absence of ASVs across different datasets using the same 18S-V9 primers, raw sequences were downloaded for the European Nucleotide Archive for Tara Oceans (PRJEB6610) (de Vargas et al. 2015), Tara Polar Circle (PRJEB9737) (Ibarbalz et al. 2019), the Malaspina Expedition (PRJEB36469) (Obiol et al. 2020), and Plankton Planet (PRJEB53961) (de Vargas et al. 2022). Raw sequences were analyzed in the QIIME2 environment as described above, except for one difference with the Tara data. As DADA2 generates an error profile based on each run, and run info for Tara samples is not available, ASVs were generated for each sample individually then combined. On average, Tara samples have more than the 100 million total bases required to learn error rates via DADA2 ensuring the accuracy of this approach (Callahan 2022). A Preston log-normal distribution was then generated from the combined datasets with the asbio package for R (Aho and Aho 2020).

To address the compositional (relative) nature of these data, two approaches were used. First, the centered log-ratio (CLR) transformation was applied, where read counts of each ASV are divided the geometric mean followed by a log transformation. This approach is a

compositional data analysis method that maintains the relationships between ASVs but results in data that are symmetric and linearly related allowing for statistical comparisons and correlations to be applied (Gloor et al. 2017; Weiss et al. 2017). Spearman correlations were calculated using ALDEx2 which converts ASV counts to probabilities via 1000 Monte Carlo instances sampled from the Dirichlet distribution that are then CLR-transformed (Fernandes et al. 2014).

Second, for all years except 2017, an internal standard of *S. pombe* genomic DNA added during cell lysis (*DNA sample collection and extraction*) to estimate ASV copies L⁻¹ as described by Lin et al. (2018). For each ASV within each sample, the number of reads was divided by the ratio of *S. pombe* reads to the number of rRNA copies added (approximately 1.29 x 10⁷ to 2.79 x 10⁷ because of differences in the amount of genomic DNA added). The total number of copies was then normalized to the volume filtered for each sample to estimate copies L⁻¹.

Rarefaction curves were generated with vegan package in R (Oksanen et al. 2007). To compare genus level abundances across the entire dataset, samples were rarefied to 24,000 reads with the QIIME2 rarefy plugin. The Shannon diversity index was calculated with the QIIME2 diversity plugin on samples with a minimum of 500 diatom reads. For beta diversity analysis, robust Aitchison distances as described by Martino et al. (2019) were calculated using the DEICODE plugin for QIIME2 on ASVs found in at least 10 samples and samples with a minimum of 500 diatom reads. PERMANOVA (adonis) and Mantel testing was then also performed with the vegan package in R (Oksanen et al. 2007).

Realized temperature niches (T_{obs}) were estimated for ASVs detected in 10 or more euphotic zone samples with the robust optimum method (Cristóbal et al. 2014). Specifically, the proportion of copies per L for each ASV were used to fill a weighted vector of a fixed size ($n = 10,000$) with corresponding temperature values (Chaffron et al. 2021). T_{obs} was then defined as

the median value of this vector. Abundance-occupancy curves and neutral model fitting was performed using the `sncm.fit` function from Burns et al. (2016). Co-occurrence networks using the euphotic zone samples including the 16S data were generated with FlashWeave v0.19.0 in sensitive mode with default options (FlashWeave-S) (Tackmann et al. 2019). Prior to network construction, ASVs whose abundances that were not in the upper quartile in at least five samples were removed (Chaffron et al. 2021). Metadata was also included during network construction to reduce spurious associations due to abiotic factors. All other statistical tests and models such as correlations, generalized additive models (GAMs), Student's T-tests, Kruskal-Wallis tests, ANOVAs, and Fisher's exact tests were performed in R. GAMs also performed with R and the `mgcv` package (Wood 2011). When performing multiple tests, P-values were adjusted with the Benjamini and Hochberg procedure (Benjamini and Hochberg 1995).

Metatranscriptome sequencing, assembly, and analysis

Poly(A) selected cDNA from total RNA was generated with the SMART-Seq® v4 Ultra® Low Input RNA Kit for Sequencing (Takara Bio USA, Inc.) which was then sheared with a Covaris® ultrasonicator. The final sequencing library was then constructed with the NEB NEBNext® Ultra™ II DNA Library Kit and sequenced on three lanes of a Novaseq 6000 with S4 flow cell (2 x 150 bp).

The resulting raw reads were then trimmed for quality and adaptor removal. Ribosomal RNA (rRNA) sequences were removed with Ribopicker v0.4.3 (Schmieder et al. 2011). Trimmed and filtered reads were then used for assembly into contigs and abundances were quantified by mapping these reads to the assembly. Both assembly and read mapping were performed with CLC Bio Genomics Server v21.0.3. Gene prediction was performed with

FragGeneScan v1.16 and rRNA removal was performed again (Rho et al. 2010). Predicted proteins were further filtered to remove those less than 10 amino acids long or with greater than or equal to 20% stop codons. Gene clusters were generated from the predicted proteins with MCL with the inflation option (-I) set to 4 and scheme option (-scheme) set to 6 (Enright et al. 2002).

With the final assembly, taxonomic annotation of each protein coding gene was assigned with the Lineage Probability Index and PhyloDB v1.076 which includes diatom transcriptomes from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) as well as some available genomes (Bertrand et al. 2015; Keeling et al. 2014; Podell and Gaasterland 2007). For functional annotation, DIAMOND BLASTP searches were performed against the Kyoto Encyclopedia of Genes and Genomes (KEGG; Release 94.1) (Bertrand et al. 2015; Kanehisa et al. 2017). KEGG Ortholog (KO) assignment was performed with KofamKOALA which utilizes hmmsearch against KOfam, an HMM database of KOs (Aramaki et al. 2019).

To normalize read counts, reads were first converted to transcripts per million (TPM) (Li et al. 2009). TPM were then used to evaluate the relative contribution of diatoms to the eukaryotic mRNA. To examine the relative expression of diatom gene clusters, diatom ORFs were selected and their TPMs were summed by gene cluster. Gene cluster expression was then expressed as the percentage of TPM relative to the total diatom TPM of each sample (Carradec et al. 2018). Absolute quantities on a transcripts L⁻¹ basis, i.e. quantitative metatranscriptomics, were also calculated based on the average recovery of the aforementioned Invitrogen™ ArrayControl™ RNA Spikes and the volume of sample filtered as described in Cohen *et al* (Cohen et al. 2022).

HPLC pigment analysis

Phytoplankton pigment concentrations were determined via high-performance liquid chromatography (HPLC). Samples were collected with 0.5, 1.04, or 2.2 L opaque brown bottles depending on the fluorescence measured by the CTD and filtered onto 25 mm GF/F filters under low vacuum pressure (≤ 40 mm Hg). Once finished, the filters were carefully folded in half, blotted on a paper towel to remove excess water, and stored in 2 mL cryovials in liquid nitrogen until analysis. Samples were then analyzed at the Horn Point Analytical Services Laboratory at the University of Maryland. Taxon-specific contributions to total chlorophyll a concentrations were determined with phytoclass v1.0.0 (Hayward et al. 2023).

Phytoplankton cell counts

From 2014 to 2020, 30 mL samples were collected at 34 CalCOFI stations with some exceptions described below (Fig. S6A). Each sample was from the second depth of the CTD-rosette (normally 10-20 m) to collect from the mixed layer when a mixed layer existed. Samples were preserved with 1% buffered formalin and returned to the lab onshore where they were pooled into the four regional samples defined by Hayward and Venrick (1998) (Fig. S7). Samples were not collected on the Winter 2014 cruise, in the Southeast Region during Winter 2018, in all regions except the Northeast during Winter 2019, and there was no cruise during spring 2020 due to the COVID-19 pandemic.

Pooled samples were counted with an inverted phase-contrast microscope. The volume of counted material depended upon the concentration of cells, which was roughly estimated by the chlorophyll concentration and varied between 0.5 mL and 100 mL. The entire sample was counted at 100X magnification for larger species and then one or more subareas were counted at

250X magnification for smaller cells. Subareas usually represented 16.7% of the total area.

When a small cell was noted on a low-power count, but was not seen under higher power, it was excluded. When possible, cells of uncertain identity were transferred to a compound scope for identification under higher magnification.

Data deposition

The sequence data reported in this study have been deposited in the National Center for Biotechnology (NCBI) sequence read archive under the BioProject accession numbers PRJNA555783, PRJNA665326, and PRJNA804265.

Acknowledgements

We thank all of the CalCOFI scientists, particularly Shonna Dovel, Megan Roadman, and Bryce Ellman as well as the captains and crews of all the research vessels involved. We also thank Mark Ohman, Kathy Barbeau, Mike Stukel for their leadership and support via the California Current Ecosystem Long-Term Ecological Research (CCE-LTER) site. Elizabeth Venrick provided all the microscopic cell counts used here, and Bryce Ellman contributed to the DNA extractions. This work was supported by the National Oceanic and Atmospheric Administration grants NA15OAR4320071 and NA19NOS4780181 (to A.E.A.), the National Science Foundation grants OCE-1026607 (to R.G.), OCE-1637632 (to R.G.), OCE-1756884 (to A.E.A.), OCE-2224726 (to A.E.A.), and DGE-1650112 (Graduate Research Fellowship to R.H.L.), the Gordon and Betty Moore Foundation grant GBMF3828 (to A.E.A.), the Simons Foundation Collaboration on Principles of Microbial Ecosystems (PriME) grant 970820 (to A.E.A.), and a Chateaubriand Fellowship from the Office of Science and Technology of the

Embassy of France in the United States and Inserm (to R.H.L). R.H.L also received partial support from the LaVerne Noyes Foundation.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Lampe, R.H., Rabines, A.J., Schulberg, A., Goericke, R., Venepally, P., Sorenson, M., Zheng, H., Semmens, B.X., Goodwin, K.D., Barton, A.D., and Allen, A.E. The dissertation author was the primary investigator and author of this paper.

References

- Abdala, Z. M., S. Clayton, S. V. Einarsson, K. Powell, C. P. Till, T. H. Coale, and P. D. Chappell. 2022. Examining ecological succession of diatoms in California Current System cyclonic mesoscale eddies. *Limnol Oceanogr* **67**: 2586-2602.
- Aho, K., and M. K. Aho. 2020. Package 'asbio'.
- Allen, W. E. 1945. Seasonal Occurrence of Marine Plankton Diatoms off Southern California in 1938. *Contributions of Scripps Institution of Oceanography, New Series*, No. 45,: 293-334.
- Amaral-Zettler, L. A., E. A. McCliment, H. W. Ducklow, and S. M. Huse. 2009. A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLoS ONE* **4**: e6372.
- Aramaki, T., R. Blanc-Mathieu, H. Endo, K. Ohkubo, M. Kanehisa, S. Goto, and H. Ogata. 2019. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**: 2251-2252.
- Armbrust, E. V. 1999. Identification of a New Gene Family Expressed during the Onset of Sexual Reproduction in the Centric Diatom *Thalassiosira weissflogii*. *Appl Environ Microbiol* **65**: 3121-3128.
- Armbrust, E. V. 2009. The life of diatoms in the world's oceans. *Nature* **459**: 185-192.
- Armstrong, F., C. Stearns, and J. Strickland. 1967. The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment, p. 381-389. *Deep Sea Research and Oceanographic Abstracts*. Elsevier.
- Barberán, A., E. O. Casamayor, and N. Fierer. 2014. The microbial contribution to macroecology. *Front Microbiol* **5**.
- Bartolek, Z., S. G. van Creveld, S. Coesel, K. R. Cain, M. Schatz, R. Morales, and E. Virginia Armbrust. 2022. Flavobacterial exudates disrupt cell cycle progression and metabolism of the diatom *Thalassiosira pseudonana*. *ISME J* **16**: 2741-2751.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* **57**: 289-300.
- Berdjeb, L., A. Parada, D. M. Needham, and J. A. Fuhrman. 2018. Short-term dynamics and interactions of marine protist communities during the spring–summer transition. *ISME J* **12**: 1907-1917.
- Bertrand, E. M., J. P. McCrow, A. Moustafa, H. Zheng, J. B. McQuaid, T. O. Delmont, A. F. Post, R. E. Sipler, J. L. Spackeen, K. Xu, D. A. Bronk, D. A. Hutchins, and A. E. Allen. 2015. Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc Natl Acad Sci USA* **112**: 9938-9943.

Biard, T., E. Bigeard, S. Audic, J. Poulain, A. Gutierrez-Rodriguez, S. Pesant, L. Stemmann, and F. Not. 2017. Biogeography and diversity of Collodaria (Radiolaria) in the global ocean. *ISME J* **11**: 1331-1344.

Bokulich, N. A., B. D. Kaehler, J. R. Rideout, M. Dillon, E. Bolyen, R. Knight, G. A. Huttley, and J. Gregory Caporaso. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**: 90.

Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. Cope, R. Da Silva, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciulek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Priesse, L. B. Rasmussen, A. Rivers, I. I. M. S. Robeson, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, and J. G. Caporaso. 2019. Reproducible, interactive, scalable, and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852-857.

Bowler, C., A. Vardi, and A. E. Allen. 2010. Oceanographic and Biogeochemical Insights from Diatom Genomes. *Annu Rev Mar Sci* **2**: 333-365.

Branscombe, L., C. Roberts, C. Widdicombe, C. Swink, W. H. Wilson, M. Cunliffe, and K. Helliwell. 2024. Seasonal niche overlap of diverse facultative antagonistic bacteria of diatoms in a productive coastal ecosystem. *bioRxiv*: 2024.2006.2005.594869.

Burns, A. R., W. Z. Stephens, K. Stagaman, S. Wong, J. F. Rawls, K. Guillemin, and B. J. M. Bohannan. 2016. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J* **10**: 655-664.

Callahan, B. 2022. A DADA2 workflow for Big Data (1.4 or later). <https://benjjneb.github.io/dada2/bigdata.html>

Callahan, B. J., P. J. McMurdie, and S. P. Holmes. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* **11**: 2639-2643.

Capone, D. G., and D. A. Hutchins. 2013. Microbial biogeochemistry of coastal upwelling regimes in a changing ocean. *Nature Geosci* **6**: 711-717.

Carpenter, J. H. 1965. The Chesapeake Bay Institute technique for the Winkler dissolved oxygen method. *Limnol Oceanogr* **10**: 141-143.

Carradec, Q., E. Pelletier, C. Da Silva, A. Alberti, Y. Seeleuthner, R. Blanc-Mathieu, G. Lima-Mendez, F. Rocha, L. Tirichine, K. Labadie, A. Kirilovsky, A. Bertrand, S. Engelen, M.-A. Madoui, R. Méheust, J. Poulain, S. Romac, D. J. Richter, G. Yoshikawa, C. Dimier, S. Kandels-Lewis, M. Picheral, S. Searson, S. G. Acinas, E. Boss, M. Follows, G. Gorsky, N. Grimsley, L. Karp-Boss, U. Krzic, S. Pesant, E. G. Reynaud, C. Sardet, M. Sieracki, S. Speich, L. Stemmann, D. Velayoudon, J. Weissenbach, O. Jaillon, J.-M. Aury, E. Karsenti, M. B. Sullivan, S. Sunagawa, P. Bork, F. Not, P. Hingamp, J. Raes, L. Guidi, H. Ogata, C. de Vargas, D. Iudicone, C. Bowler, P. Wincker, and C. Tara Oceans. 2018. A global ocean atlas of eukaryotic genes. *Nat Commun* **9**: 373.

Cermeño, P., and P. G. Falkowski. 2009. Controls on diatom biogeography in the ocean. *Science* **325**: 1539-1541.

Chaffron, S., E. Delage, M. Budinich, D. Vintache, N. Henry, C. Nef, M. Ardyna, A. A. Zayed, P. C. Junger, and P. E. Galand. 2021. Environmental vulnerability of the global ocean epipelagic plankton community interactome. *Sci Adv* **7**: eabg1921.

Chavez, F. P., and M. Messié. 2009. A comparison of Eastern Boundary Upwelling Ecosystems. *Prog Oceanogr* **83**: 80-96.

Checkley, D. M., and J. A. Barth. 2009. Patterns and processes in the California Current System. *Prog Oceanogr* **83**: 49-64.

Cohen, N. R., H. Alexander, A. I. Krinos, S. K. Hu, and R. H. Lampe. 2022. Marine Microeukaryote Metatranscriptomics: Sample Processing and Bioinformatic Workflow Recommendations for Ecological Applications. *Front Mar Sci* **9**: 867007.

Cohen, N. R., K. A. Ellis, R. H. Lampe, H. M. McNair, B. S. Twining, M. A. Brzezinski, M. T. Maldonado, F. I. Kuzminov, K. Thamatrakoln, C. P. Till, K. W. Bruland, W. G. Sunda, S. Bargu, and A. Marchetti. 2017. Variations in diatom transcriptional responses to changes in iron availability across ocean provinces. *Front Mar Sci* **4**: 360.

Cristóbal, E., S. V. Ayuso, A. Justel, and M. Toro. 2014. Robust optima and tolerance ranges of biological indicators: a new method to identify sentinels of global warming. *Ecological Research* **29**: 55-68.

de Vargas, C., S. Audic, N. Henry, J. Decelle, F. Mahé, R. Logares, E. Lara, C. Berney, N. Le Bescot, I. Probert, M. Carmichael, J. Poulain, S. Romac, S. Colin, J.-M. Aury, L. Bittner, S. Chaffron, M. Dunthorn, S. Engelen, O. Flegontova, L. Guidi, A. Horák, O. Jaillon, G. Lima-Mendez, J. Lukeš, S. Malviya, R. Morard, M. Mulot, E. Scalco, R. Siano, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, S. G. Acinas, P. Bork, C. Bowler, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, F. Not, H. Ogata, S. Pesant, J. Raes, M. E. Sieracki, S. Speich, L. Stemmann, S. Sunagawa, J. Weissenbach, P. Wincker, and E. Karsenti. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**: 1261605.

de Vargas, C., N. Le Bescot, T. Pollina, N. Henry, S. Romac, S. Colin, N. Haëntjens, M. Carmichael, C. Berger, D. Le Guen, J. Decelle, F. Mahé, J. Poulain, E. Malpot, C. Beaumont, M. Hardy, D. Guiffant, I. Probert, D. F. Gruber, A. E. Allen, G. Gorsky, M. J. Follows, X. Pochon,

- R. Troublé, B. B. Cael, F. Lombard, E. Boss, M. Prakash, t. P. P. c. t. , R. Bazile, E. Boss, G. Bourdin, B. Cael, R. Casati, S. b. Colin, C. d. Vargas, G. Gorsky, D. Guiffant, N. Haentjens, N. Henry, A. Larson, N. L. Bescot, F. Lombard, G. Mirambeau, C. m. Moulin, A. Oddone, M. Prakash, C. Prazuck, V. Raimbault, C. Trellu, and R. Troublé. 2022. Plankton Planet: A frugal, cooperative measure of aquatic life at the planetary scale. *Front Mar Sci* **9**.
- Deutsch, C., H. Frenzel, J. C. McWilliams, L. Renault, F. Kessouri, E. Howard, J.-H. Liang, D. Bianchi, and S. Yang. 2021. Biogeochemical variability in the California Current System. *Prog Oceanogr* **196**: 102565.
- Di Lorenzo, E., and N. Mantua. 2016. Multi-year persistence of the 2014/15 North Pacific marine heatwave. *Nat Clim Change* **6**: 1042-1047.
- Di Lorenzo, E., N. Schneider, K. M. Cobb, P. J. S. Franks, K. Chhak, A. J. Miller, J. C. McWilliams, S. J. Bograd, H. Arango, E. Curchitser, T. M. Powell, and P. Rivière. 2008. North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophys Res Letters* **35**, L08607.
- Durkin, C. A., I. Cetinić, M. Estapa, Z. Ljubešić, M. Mucko, A. Neeley, and M. Omand. 2022. Tracing the path of carbon export in the ocean through DNA sequencing of individual sinking particles. *ISME J* **16**: 1896-1906.
- Enright, A. J., S. Van Dongen, and C. A. Ouzounis. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* **30**: 1575-1584.
- Eppley, R. W. 1992. Chlorophyll, photosynthesis and new production in the Southern California Bight. *Prog Oceanogr* **30**: 117-150.
- Estrada, M., and D. Blasco. 1985. Phytoplankton assemblages in coastal upwelling areas, p. 379-402. *In* C. Bas, R. Margalef and P. Rubies [eds.], *International Symposium on the Most Important Upwelling Areas off Western Africa*.
- Fernandes, A. D., J. N. S. Reid, J. M. Macklaim, T. A. McMurrough, D. R. Edgell, and G. B. Gloor. 2014. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome* **2**: 15.
- Fiechter, J., and A. M. Moore. 2024. Physical and Biogeochemical Properties of California Current Upwelled Source Waters. *J Geophys Res Oceans* **129**: e2023JC020164.
- Forsch, K. O., K. C. Fulton, M. M. Weiss, J. W. Krause, M. R. Stukel, and K. A. Barbeau. 2023. Iron limitation and biogeochemical effects in southern California Current coastal upwelling filaments. *J Geophys Res Oceans* **128**: e2023JC019961.
- Gloor, G. B., J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. 2017. Microbiome datasets are compositional: and this is not optional. *Front Microbiol* **8**: 2224.

- Goericke, R. 2011. The structure of marine phytoplankton communities-patterns, rules, and mechanisms. California Cooperative Oceanic Fisheries Investigations Report **52**: 182-197.
- Gordon, L., J. Jennings, A. Ross, and J. Krest. 1992. A suggested protocol for continuous flow automated analysis of seawater nutrients in the WOCE hydrographic program and the Joint Global Ocean Fluxes Study. Grp. Tech Rpt: 92-91.
- Gross, M. 2012. The mysteries of the diatoms. *Curr Biol* **22**: R581-R585.
- Guillou, L., D. Bachar, S. Audic, D. Bass, C. Berney, L. Bittner, C. Boute, G. Burgaud, C. de Vargas, J. Decelle, J. del Campo, J. R. Dolan, M. Dunthorn, B. Edvardsen, M. Holzmann, W. H. C. F. Kooistra, E. Lara, N. Le Bescot, R. Logares, F. Mahé, R. Massana, M. Montresor, R. Morard, F. Not, J. Pawlowski, I. Probert, A.-L. Sauvadet, R. Siano, T. Stoeck, D. Vaultot, P. Zimmermann, and R. Christen. 2012. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* **41**: D597-D604.
- Guiry, M. D. 2024. How many species of algae are there? A reprise. Four kingdoms, 14 phyla, 63 classes and still growing. *J Phycol* **60**: 214-228.
- Hasle, G. R., and E. E. Syvertsen. 1997. Chapter 2 - Marine Diatoms, p. 5-385. *In* C. R. Tomas [ed.], *Identifying Marine Phytoplankton*. Academic Press.
- Hayward, A., M. H. Pinkerton, and A. Gutierrez-Rodriguez. 2023. phytoclass: A pigment-based chemotaxonomic method to determine the biomass of phytoplankton classes. *Limnol Oceanogr-Meth* **21**: 220-241.
- Hayward, T. L., and E. L. Venrick. 1998. Nearsurface pattern in the California Current: coupling between physical and biological structure. *Deep-Sea Res II: Top Stud Oceanogr* **45**: 1617-1638.
- Hill, M. O. 1973. Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology* **54**: 427-432.
- Hogle, S. L., C. L. Dupont, B. M. Hopkinson, A. L. King, K. N. Buck, K. L. Roe, R. K. Stuart, A. E. Allen, E. L. Mann, Z. I. Johnson, and K. A. Barbeau. 2018. Pervasive iron limitation at subsurface chlorophyll maxima of the California Current. *Proc Nat Acad Sci USA* **115**: 13300-13305.
- Hutchins, D. A., and K. W. Bruland. 1998. Iron-limited diatom growth and Si:N uptake ratios in a coastal upwelling regime. *Nature* **393**: 561-564.
- Ibarbalz, F. M., N. Henry, M. C. Brandão, S. Martini, G. Busseni, H. Byrne, L. P. Coelho, H. Endo, J. M. Gasol, A. C. Gregory, F. Mahé, J. Rignonato, M. Royo-Llonch, G. Salazar, I. Sanz-Sáez, E. Scalco, D. Siviadan, A. A. Zayed, A. Zingone, K. Labadie, J. Ferland, C. Marec, S. Kandels, M. Picheral, C. Dimier, J. Poulain, S. Pisarev, M. Carmichael, S. Pesant, S. G. Acinas, M. Babin, P. Bork, E. Boss, C. Bowler, G. Cochrane, C. de Vargas, M. Follows, G. Gorsky, N. Grimsley, L. Guidi, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels, L. Karp-Boss, E. Karsenti, F. Not, H. Ogata, S. Pesant, N. Poulton, J. Raes, C. Sardet, S. Speich, L. Stemmann, M. B.

Sullivan, S. Sunagawa, P. Wincker, M. Babin, E. Boss, D. Iudicone, O. Jaillon, S. G. Acinas, H. Ogata, E. Pelletier, L. Stemmann, M. B. Sullivan, S. Sunagawa, L. Bopp, C. de Vargas, L. Karp-Boss, P. Wincker, F. Lombard, C. Bowler, and L. Zinger. 2019. Global Trends in Marine Plankton Diversity across Kingdoms of Life. *Cell* **179**: 1084-1097.e1021.

Jabbur, M. L., B. P. Bratton, and C. H. Johnson. 2024. Bacteria can anticipate the seasons: Photoperiodism in cyanobacteria. *Science* **385**: 1105-1111.

Jacox, M. G., M. A. Alexander, N. J. Mantua, J. D. Scott, G. Hervieux, R. S. Webb, and F. E. Werner. 2018a. Forcing of multi-year extreme ocean temperatures that impacted California Current living marine resources in 2016. *Bull. Amer. Meteor. Soc* **99**.

Jacox, M. G., C. A. Edwards, E. L. Hazen, and S. J. Bograd. 2018b. Coastal upwelling revisited: Ekman, Bakun, and improved upwelling indices for the US West Coast. *J Geophys Res Oceans* **123**: 7332-7350.

James, C. C., A. D. Barton, L. Z. Allen, R. H. Lampe, A. Rabines, A. Schulberg, H. Zheng, R. Goericke, K. D. Goodwin, and A. E. Allen. 2022. Influence of nutrient supply on plankton microbiome biodiversity and distribution in a coastal upwelling region. *Nat Commun* **13**: 2448.

Kanehisa, M., M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353-D361.

Keeling, P. J., F. Burki, H. M. Wilcox, B. Allam, E. E. Allen, L. A. Amaral-Zettler, E. V. Armbrust, J. M. Archibald, A. K. Bharti, C. J. Bell, B. Beszteri, K. D. Bidle, C. T. Cameron, L. Campbell, D. A. Caron, R. A. Cattolico, J. L. Collier, K. Coyne, S. K. Davy, P. Deschamps, S. T. Dyhrman, B. Edvardsen, R. D. Gates, C. J. Gobler, S. J. Greenwood, S. M. Guida, J. L. Jacobi, K. S. Jakobsen, E. R. James, B. Jenkins, U. John, M. D. Johnson, A. R. Juhl, A. Kamp, L. A. Katz, R. Kiene, A. Kudryavtsev, B. S. Leander, S. Lin, C. Lovejoy, D. Lynn, A. Marchetti, G. McManus, A. M. Nedelcu, S. Menden-Deuer, C. Miceli, T. Mock, M. Montresor, M. A. Moran, S. Murray, G. Nadathur, S. Nagai, P. B. Ngam, B. Palenik, J. Pawlowski, G. Petroni, G. Piganeau, M. C. Posewitz, K. Rengefors, G. Romano, M. E. Rumpho, T. Rynearson, K. B. Schilling, D. C. Schroeder, A. G. B. Simpson, C. H. Slamovits, D. R. Smith, G. J. Smith, S. R. Smith, H. M. Sosik, P. Stief, E. Theriot, S. N. Twary, P. E. Umale, D. Vaultot, B. Wawrik, G. L. Wheeler, W. H. Wilson, Y. Xu, A. Zingone, and A. Z. Worden. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLOS Biol* **12**: e1001889.

Kemp, A. E. S., and T. A. Villareal. 2018. The case of the diatoms and the muddled mandalas: Time to recognize diatom adaptations to stratified waters. *Prog Oceanogr* **167**: 138-149.

King, A. L., and K. A. Barbeau. 2011. Dissolved iron and macronutrient distributions in the southern California Current System. *J Geophys Res Oceans* **116**.

Kolody, B. C., J. P. McCrow, L. Z. Allen, F. O. Aylward, K. M. Fontanez, A. Moustafa, M. Moniruzzaman, F. P. Chavez, C. A. Scholin, E. E. Allen, A. Z. Worden, E. F. Delong, and A. E. Allen. 2019. Diel transcriptional response of a California Current plankton microbiome to light, low iron, and enduring viral infection. *ISME J* **13**: 2817-2833.

Kolody, B. C., S. R. Smith, L. Z. Allen, J. P. McCrow, A. Moustafa, D. Shi, B. M. Hopkinson, F. M. M. Morel, B. B. Ward, and A. E. Allen. 2022. Nitrogen and Iron Availability Drive Metabolic Remodeling and Natural Selection of Diverse Phytoplankton during Experimental Upwelling. *mSystems* **7**: e00729-00722.

Kooistra, W. H. C. F., R. Gersonde, L. K. Medlin, and D. G. Mann. 2007. CHAPTER 11 - The Origin and Evolution of the Diatoms: Their Adaptation to a Planktonic Existence, p. 207-249. *In* P. G. Falkowski and A. H. Knoll [eds.], *Evolution of Primary Producers in the Sea*. Academic Press.

Krause, J. W., M. A. Brzezinski, R. Goericke, M. R. Landry, M. D. Ohman, M. R. Stukel, and A. G. Taylor. 2015. Variability in diatom contributions to biomass, organic matter production and export across a frontal gradient in the California Current Ecosystem. *J Geophys Res Oceans* **120**: 1032-1047.

Kröger, N., and N. Poulsen. 2008. Diatoms—From Cell Wall Biogenesis to Nanotechnology. *Annu Rev Genet* **42**: 83-107.

Kuhlisch, C., A. Shemi, N. Barak-Gavish, D. Schatz, and A. Vardi. 2024. Algal blooms in the ocean: hot spots for chemically mediated microbial interactions. *Nat Rev Micro* **22**: 138-154.

Lampe, R., A. Rabines, S. Wood, A. Schulberg, R. Goericke, P. Venepally, H. Zheng, M. Stukel, M. Landry, A. Barton, and A. Allen. 2024a. Relationships between phytoplankton pigments and DNA- or RNA-based abundances support ecological applications. *EGUsphere*: <https://doi.org/10.5194/egusphere-2024-3285>.

Lampe, R. H., T. H. Coale, K. O. Forsch, L. J. Jabre, S. Kekuewa, E. M. Bertrand, A. Horák, M. Oborník, A. J. Rabines, E. Rowland, H. Zheng, A. J. Andersson, K. A. Barbeau, and A. E. Allen. 2023. Short-term acidification promotes diverse iron acquisition and conservation mechanisms in upwelling-associated phytoplankton. *Nat Commun* **14**: 7215.

Lampe, R. H., T. H. Coale, J. B. McQuaid, and A. E. Allen. 2024b. Molecular Mechanisms for Iron Uptake and Homeostasis in Marine Eukaryotic Phytoplankton. *Ann Rev Micro* **78**.

Lampe, R. H., N. R. Cohen, K. A. Ellis, K. W. Bruland, M. T. Maldonado, T. D. Peterson, C. E. Parker, S. Bargu, M. A. Brzezinski, K. Thamtracoln, F. I. Kuzminov, B. S. Twining, and A. Marchetti. 2018a. Divergent gene expression among phytoplankton taxa in response to upwelling. *Environ Microbiol* **20**: 3069-3082.

Lampe, R. H., E. L. Mann, N. R. Cohen, C. P. Till, K. Thamtracoln, M. A. Brzezinski, K. W. Bruland, B. S. Twining, and A. Marchetti. 2018b. Different iron storage strategies among bloom-forming diatoms. *Proc Natl Acad Sci USA* **115**: E12275-E12284.

Leblanc, K., B. Queguiner, F. Diaz, V. Cornet, M. Michel-Rodriguez, X. Durrieu de Madron, C. Bowler, S. Malviya, M. Thyssen, and G. Grégori. 2018. Nanoplanktonic diatoms are globally overlooked but play a role in spring blooms and carbon export. *Nat Commun* **9**: 953.

- Leising, A. W., I. D. Schroeder, S. J. Bograd, J. Abell, R. Durazo, G. Gaxiola-Castro, E. P. Bjorkstedt, J. Field, K. Sakuma, and R. R. Robertson. 2015. State of the California Current 2014-15: Impacts of the Warm-Water" Blob". California Cooperative Oceanic Fisheries Investigations Reports **56**.
- Li, B., V. Ruotti, R. M. Stewart, J. A. Thomson, and C. N. Dewey. 2009. RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics* **26**: 493-500.
- Lin, Y., S. Gifford, H. Ducklow, O. Schofield, and N. Cassar. 2018. Towards quantitative microbiome community profiling using internal standards. *Appl Environ Microbiol*: AEM.02634-02618.
- Lin, Y., O. Torano, L. Whitehouse, E. Pierce, C. P. Till, M. Hurst, R. Freiburger, T. Mellett, M. T. Maldonado, J. Guo, M. Sutton, D. Zeitz, and A. Marchetti. 2024. Variability in the phytoplankton response to upwelling across an iron limitation mosaic within the California current system. *Limnol Oceanogr* **69**: 888-901.
- Malviya, S., E. Scalco, S. Audic, F. Vincent, A. Veluchamy, J. Poulain, P. Wincker, D. Iudicone, C. de Vargas, L. Bittner, A. Zingone, and C. Bowler. 2016. Insights into global diatom distribution and diversity in the world's ocean. *Proc Natl Acad Sci USA* **113**: E1516-E1525.
- Maniscalco, M. A., M. A. Brzezinski, J. W. Krause, and K. Thamatrakoln. 2023. Decoupling silicon metabolism from carbon and nitrogen assimilation poises diatoms to exploit episodic nutrient pulses in a coastal upwelling system. *Front Mar Sci* **10**: 1291294.
- Mann, D. G., and S. Droop. 1996. Biodiversity, biogeography and conservation of diatoms, p. 19-32. *Biogeography of Freshwater Algae: Proceedings of the Workshop on Biogeography of Freshwater Algae, Held During the Fifth International Phycological Congress, Qingdao, China, June 1994*. Springer.
- Margalef, R. 1978. Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanol acta* **1**: 493-509.
- Martin, J. L., I. Santi, P. Pitta, U. John, and N. Gypens. 2022. Towards quantitative metabarcoding of eukaryotic plankton: an approach to improve 18S rRNA gene copy number bias. *Metabarcoding Metagenomics* **6**.
- Martin, K., K. Schmidt, A. Toseland, C. A. Boulton, K. Barry, B. Beszteri, C. P. D. Brussaard, A. Clum, C. G. Daum, E. Elie-Fadrosh, A. Fong, B. Foster, B. Foster, M. Ginzburg, M. Huntemann, N. N. Ivanova, N. C. Kyrpides, E. Lindquist, S. Mukherjee, K. Palaniappan, T. B. K. Reddy, M. R. Rizkallah, S. Roux, K. Timmermans, S. G. Tringe, W. H. van de Poll, N. Varghese, K. U. Valentin, T. M. Lenton, I. V. Grigoriev, R. M. Leggett, V. Moulton, and T. Mock. 2021. The biogeographic differentiation of algal microbiomes in the upper ocean from pole to pole. *Nat Commun* **12**: 5483.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 10-12.

- Martino, C., J. T. Morton, C. A. Marotz, L. R. Thompson, A. Tripathi, R. Knight, and K. Zengler. 2019. A Novel Sparse Compositional Technique Reveals Microbial Perturbations. *mSystems* **4**: e00016-00019.
- Matthews, S. A., and M. D. Ohman. 2023. A space-for-time framework for forecasting the effects of ocean stratification on zooplankton vertical habitat use and trait composition. *Limnol Oceanogr* **68**: 2688-2702.
- McNichol, J., P. M. Berube, S. J. Biller, and J. A. Fuhrman. 2021. Evaluating and Improving SSU rRNA PCR Primer Coverage for Bacteria, Archaea, and Eukaryotes Using Metagenomes from Global Ocean Surveys. *bioRxiv*: 2020.2011.2009.375543.
- McQuaid, J. B., A. B. Kustka, M. Oborník, A. Horák, J. P. McCrow, B. J. Karas, H. Zheng, T. Kindeberg, A. J. Andersson, K. A. Barbeau, and A. E. Allen. 2018. Carbonate-sensitive phytoferritin controls high-affinity iron uptake in diatoms. *Nature* **555**: 534.
- Murphy, J., and J. P. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. *Anal Chim Acta* **27**: 31-36.
- Obiol, A., C. R. Giner, P. Sánchez, C. M. Duarte, S. G. Acinas, and R. Massana. 2020. A metagenomic assessment of microbial eukaryotic diversity in the global ocean. *Mol Ecol Resour* **20**: 718-731.
- Oksanen, J., R. Kindt, P. Legendre, B. O'Hara, M. H. H. Stevens, M. J. Oksanen, and M. Suggets. 2007. The vegan package. *Community ecology package* **10**: 719.
- Parada, A. E., D. M. Needham, and J. A. Fuhrman. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* **18**: 1403-1414.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg. 2011. Scikit-learn: Machine learning in Python. *J Mach Learn Res* **12**: 2825-2830.
- Pierella Karlusich, J. J., K. Cosnier, L. Zinger, N. Henry, C. Nef, G. Bernard, E. Scalco, E. Dvorak, F. R. Jimenez Vieira, E. Delage, S. Chaffron, S. Ovchinnikov, A. Zingone, and C. Bowler. 2024. Patterns and drivers of diatom diversity and abundance in the global ocean. *bioRxiv*: 2024.2006.2008.598090.
- Podell, S., and T. Gaasterland. 2007. DarkHorse: a method for genome-wide prediction of horizontal gene transfer. *Genome Biol* **8**: R16.
- Preston, C. M., C. A. Durkin, and K. M. Yamahara. 2020. DNA metabarcoding reveals organisms contributing to particulate matter flux to abyssal depths in the North East Pacific ocean. *Deep-Sea Res II: Top Stud Oceanogr* **173**: 104708.
- Rho, M., H. Tang, and Y. Ye. 2010. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* **38**: e191.

Röttjers, L., and K. Faust. 2018. From hairballs to hypotheses—biological insights from microbial networks. *FEMS Microbiol Rev* **42**: 761-780.

Rynearson, T. A., S. A. Flickinger, and D. N. Fontaine. 2020. Metabarcoding Reveals Temporal Patterns of Community Composition and Realized Thermal Niches of *Thalassiosira* spp. (Bacillariophyceae) from the Narragansett Bay Long-Term Plankton Time Series. *Biology* **9**: 19.

Schmieder, R., Y. W. Lim, and R. Edwards. 2011. Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* **28**: 433-435.

Sekula-Wood, E., A. Schnetzer, C. R. Benitez-Nelson, C. Anderson, W. M. Berelson, M. A. Brzezinski, J. M. Burns, D. A. Caron, I. Cetinic, and J. L. Ferry. 2009. Rapid downward transport of the neurotoxin domoic acid in coastal waters. *Nat Geosci* **2**: 272-275.

Setta, S. P., S. Lerch, B. D. Jenkins, S. T. Dyhrman, and T. A. Rynearson. 2023. Oligotrophic waters of the Northwest Atlantic support taxonomically diverse diatom communities that are distinct from coastal waters. *J Phycol* **59**: 1202-1216.

Shade, A., and N. Stopnisek. 2019. Abundance-occupancy distributions to prioritize plant core microbiome membership. *Curr Opin Microbiol* **49**: 50-58.

Sloan, W. T., M. Lunn, S. Woodcock, I. M. Head, S. Nee, and T. P. Curtis. 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* **8**: 732-740.

Stoecker, D. K., P. J. Hansen, D. A. Caron, and A. Mitra. 2017. Mixotrophy in the marine plankton. *Annu Rev Mar Sci* **9**: 311-335.

Strickland, J. D. H., and T. R. Parsons. 1972. A practical handbook of seawater analysis. Fisheries Research Board of Canada.

Stukel, M. R., J. P. Irving, T. B. Kelly, M. D. Ohman, C. K. Fender, and N. Yingling. 2023. Carbon sequestration by multiple biological pump pathways in a coastal upwelling biome. *Nat Commun* **14**: 2024.

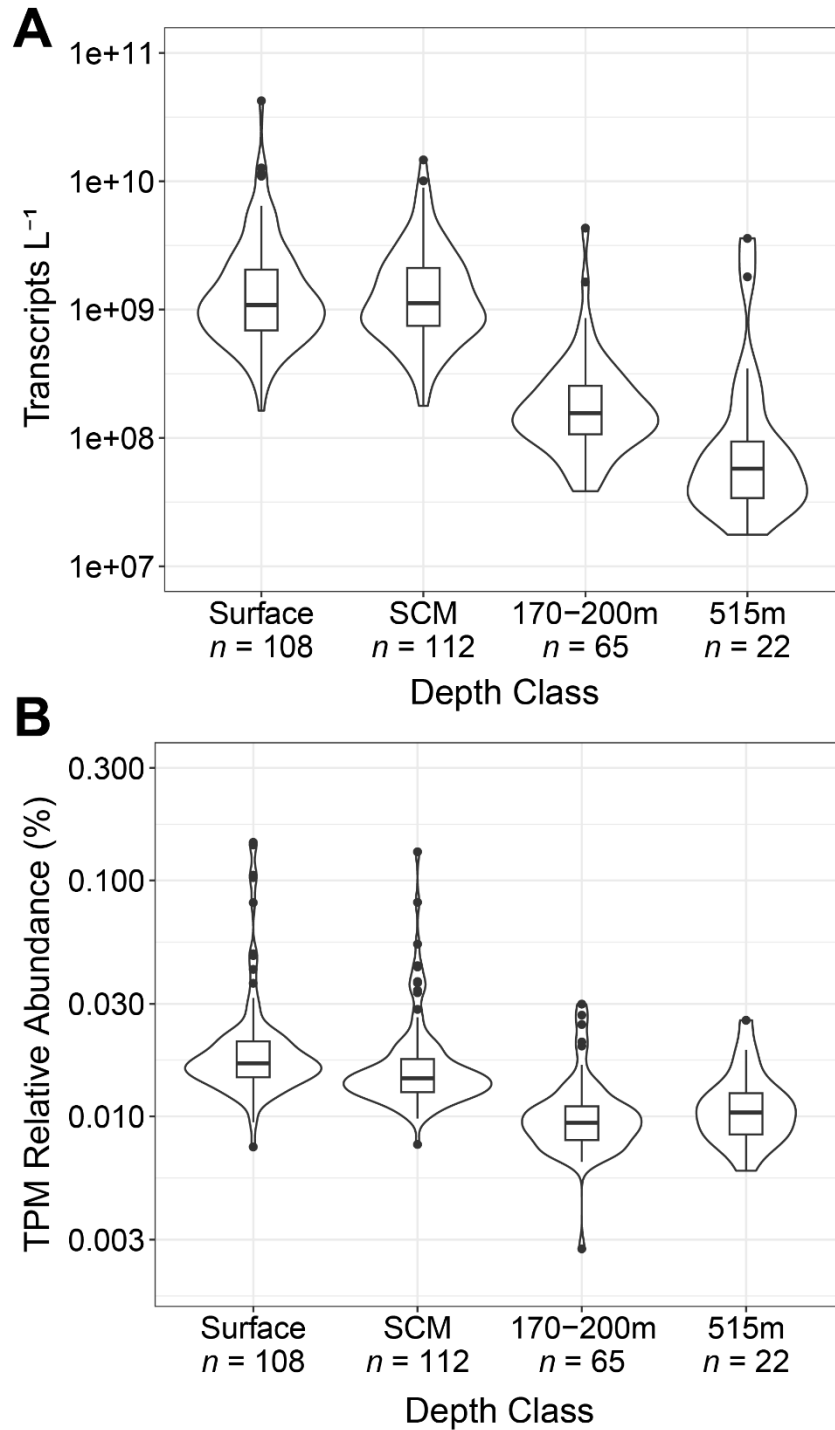
Sunda, W. G., and A. Marchetti. 2024. Proton-pumping rhodopsins promote the growth and survival of phytoplankton in a highly variable ocean. *The ISME Journal* **18**.

Tackmann, J., J. F. M. Rodrigues, and C. von Mering. 2019. Rapid inference of direct interactions in large-scale ecological networks from heterogeneous microbial sequencing data. *Cell Systems* **9**: 286-296. e288.

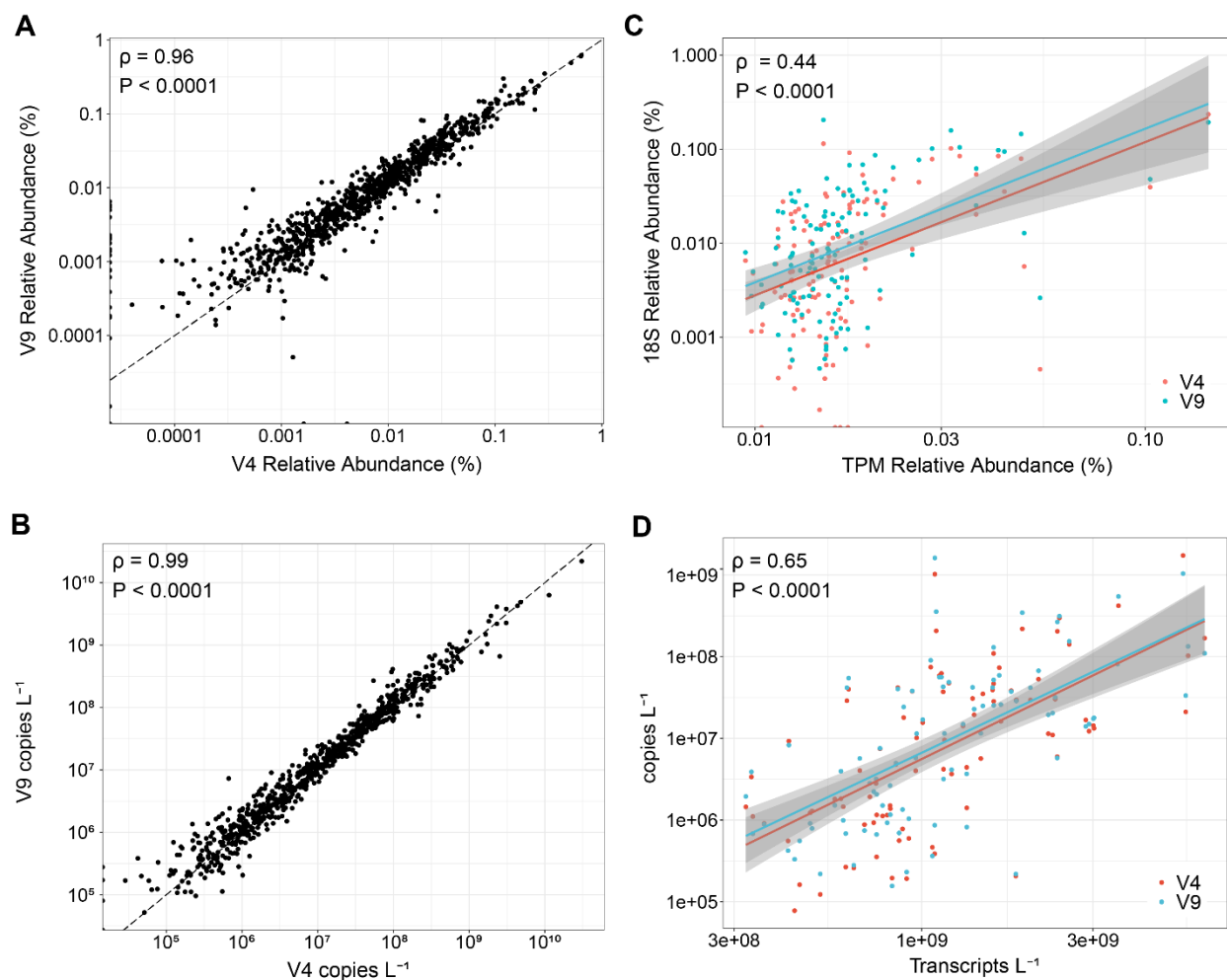
Taylor, A. G., and M. R. Landry. 2018. Phytoplankton biomass and size structure across trophic gradients in the southern California Current and adjacent ocean ecosystems. *Mar Ecol Prog Ser* **592**: 1-17.

- Valencia, B., M. R. Stukel, A. E. Allen, J. P. McCrow, A. Rabines, and M. R. Landry. 2022. Microbial communities associated with sinking particles across an environmental gradient from coastal upwelling to the oligotrophic ocean. *Deep Sea Res Part I Oceanogr Res Pap* **179**: 103668.
- van Tol, H. M., S. A. Amin, and E. V. Armbrust. 2016. Ubiquitous marine bacterium inhibits diatom cell division. *ISME J* **11**: 31-42.
- Vanormelingen, P., E. Verleyen, W. Vyverman. 2007. The diversity and distribution of diatoms: from cosmopolitanism to narrow endemisms, p. 159-171. In W. Foissner and D.L. Hawksworth [eds.], *Protist Diversity and Geographical Distribution. Topics in Biodiversity and Conservation*, vol 8. Springer.
- Venrick, E. 1998. Spring in the California Current: the distribution of phytoplankton species, April 1993 and April 1995. *Mar Ecol Prog Ser* **167**: 73-88.
- Venrick, E. 2002. Floral patterns in the California Current System off southern California: 1990-1996. *J Mar Res* **60**: 171-189.
- Venrick, E. 2009. Floral patterns in the California Current: The coastal-offshore boundary zone. *J Mar Res* **67**: 89-111.
- Venrick, E. L. 2012. Phytoplankton in the California Current system off southern California: Changes in a changing environment. *Prog Oceanogr* **104**: 46-58.
- Weiss, S., Z. Z. Xu, S. Peddada, A. Amir, K. Bittinger, A. Gonzalez, C. Lozupone, J. R. Zaneveld, Y. Vázquez-Baeza, A. Birmingham, E. R. Hyde, and R. Knight. 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5**: 27.
- Werb, B. E., and D. L. Rudnick. 2023. Remarkable Changes in the Dominant Modes of North Pacific Sea Surface Temperature. *Geophys Res Lett* **50**: e2022GL101078.
- Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**: 3-36.
- Zhu, F., R. Massana, F. Not, D. Marie, and D. Vaultot. 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol* **52**: 79-92.
- Zimmermann, J., G. Glöckner, R. Jahn, N. Enke, and B. Gemeinholzer. 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Mol Ecol Resour* **15**: 526-542.

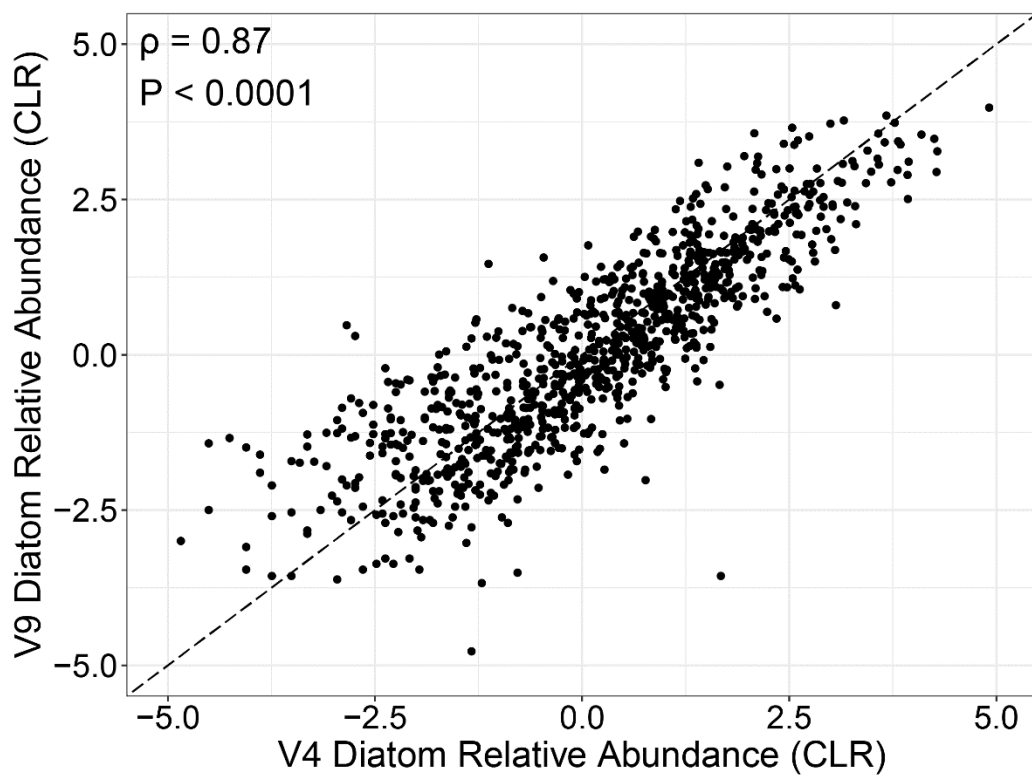
CHAPTER 4 SUPPLEMENTAL INFORMATION



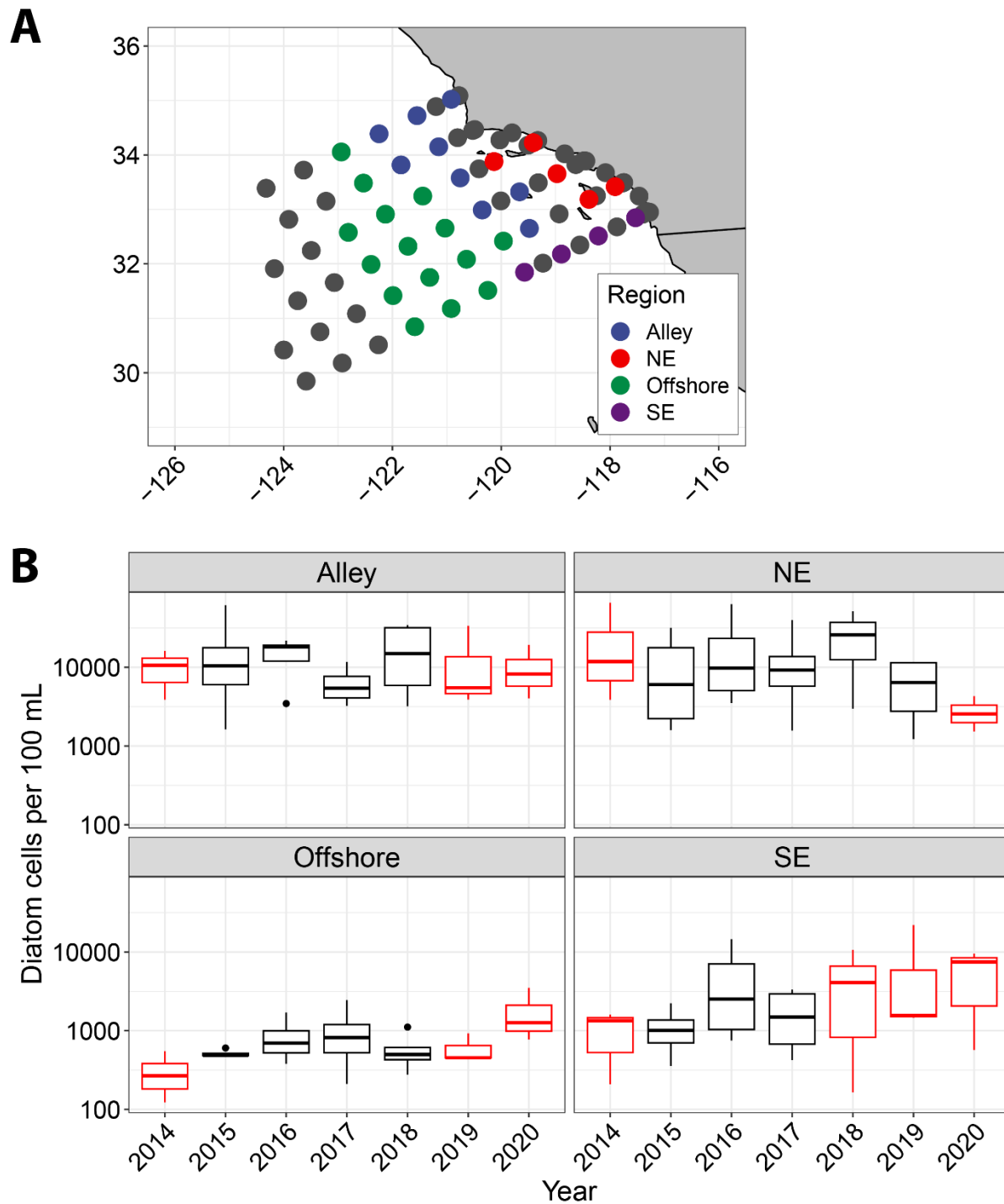
Supplemental Figure 4.1: The abundance of diatom transcripts as (A) absolute abundances (transcripts L^{-1}) and (B) relative abundances (percent of TPM) for each depth category. The number of samples is denoted as n .



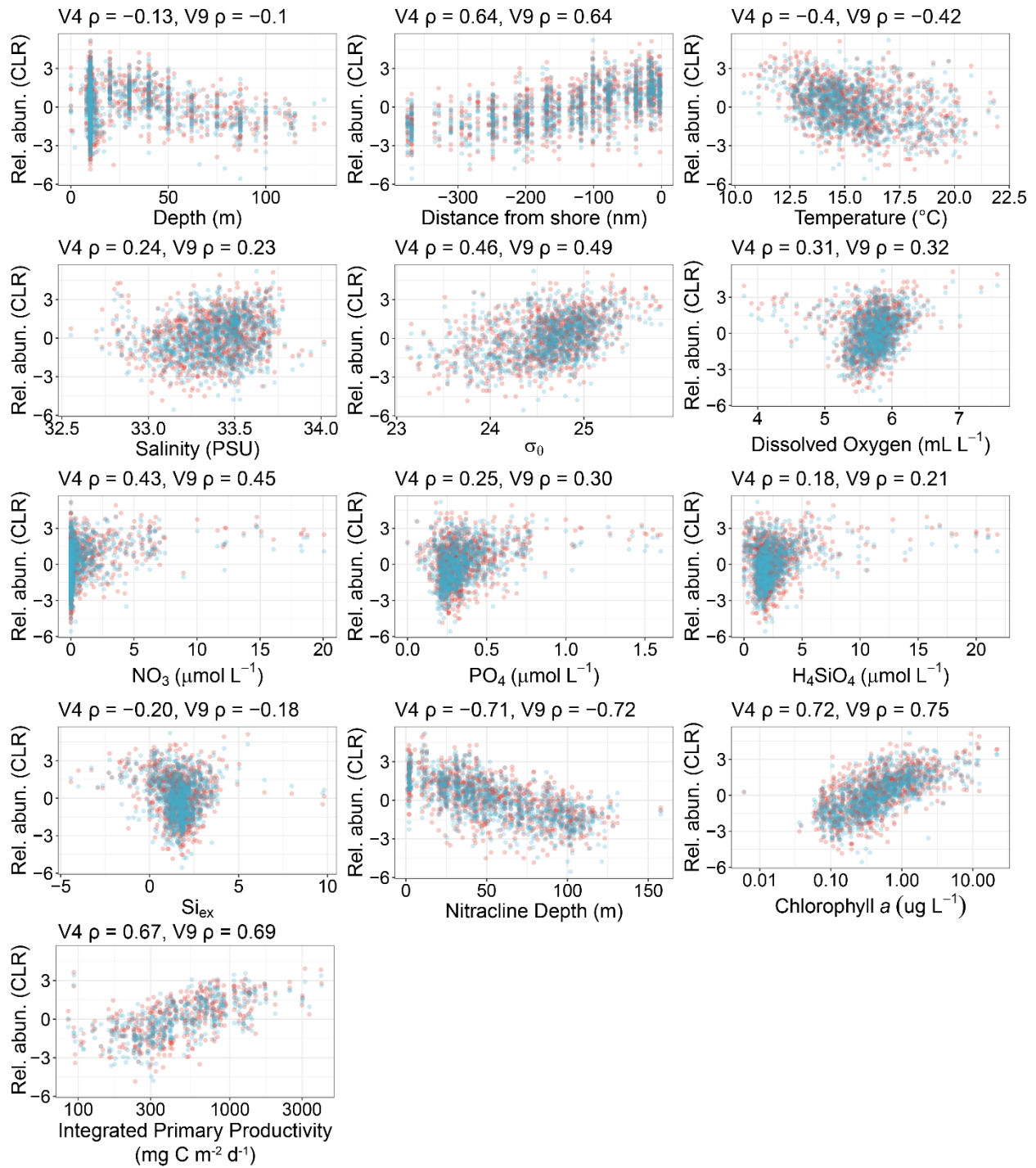
Supplemental Figure 4.2: Spearman correlations between different measures of diatom abundances. (A, B) Relative and absolute abundances for V4 versus V9. The dashed lines indicate a one-to-one relationship. (C) Relative transcript abundances versus relative 18S abundances for V4 (red) and (V9) blue. (D) Absolute transcript abundances versus absolute 18S abundances for V4 (red) and (V9) blue. In panels C and D, lines show linear models with 95% confidence intervals.



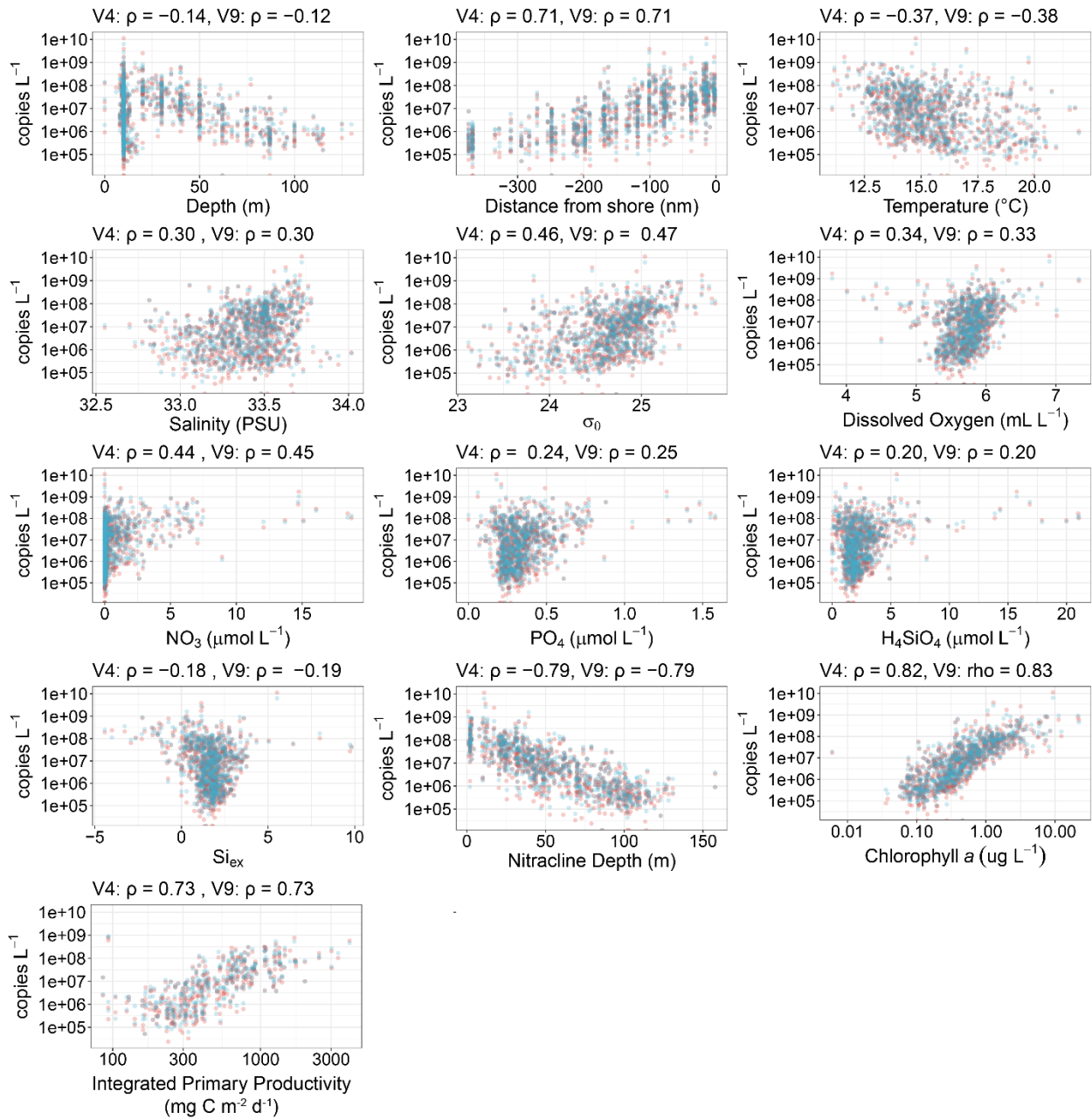
Supplemental Figure 4.3: 18S-V4 CLR-transformed relative abundances versus 18S-V9 CLR-transformed relative abundances for summed diatom ASVs in each NCOG sample. The dashed line indicates a one-to-one relationship. Relative abundances are significantly correlated (Spearman's $\rho = 0.87$, $P < 0.0001$).



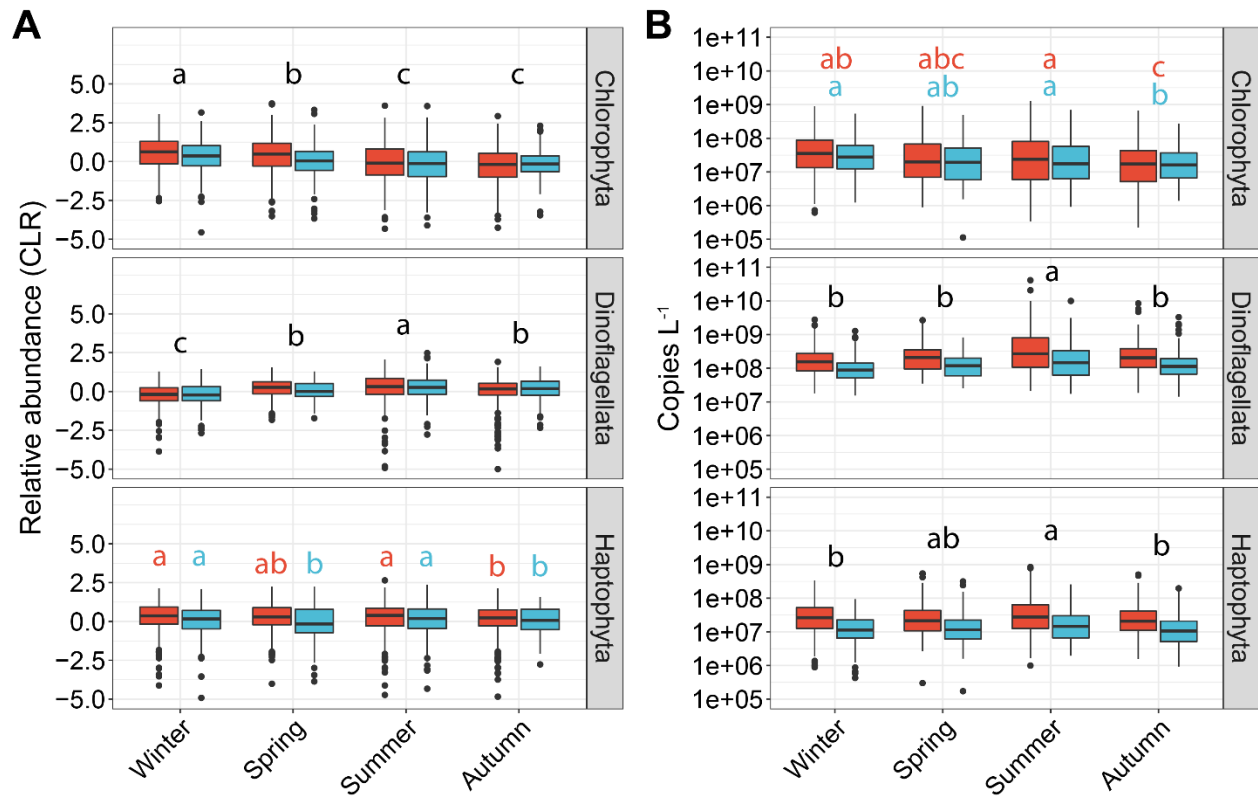
Supplemental Figure 4.4: Diatom cell counts from near-surface waters and pooled by regions on CalCOFI cruises from 2014 to 2020. (A) Stations coded by region for pooling the cell count samples. (B) Total diatom cells per 100 mL for each region and year. Regions and years that are highlighted in red do not have complete sample sets (see *Materials and Methods*).



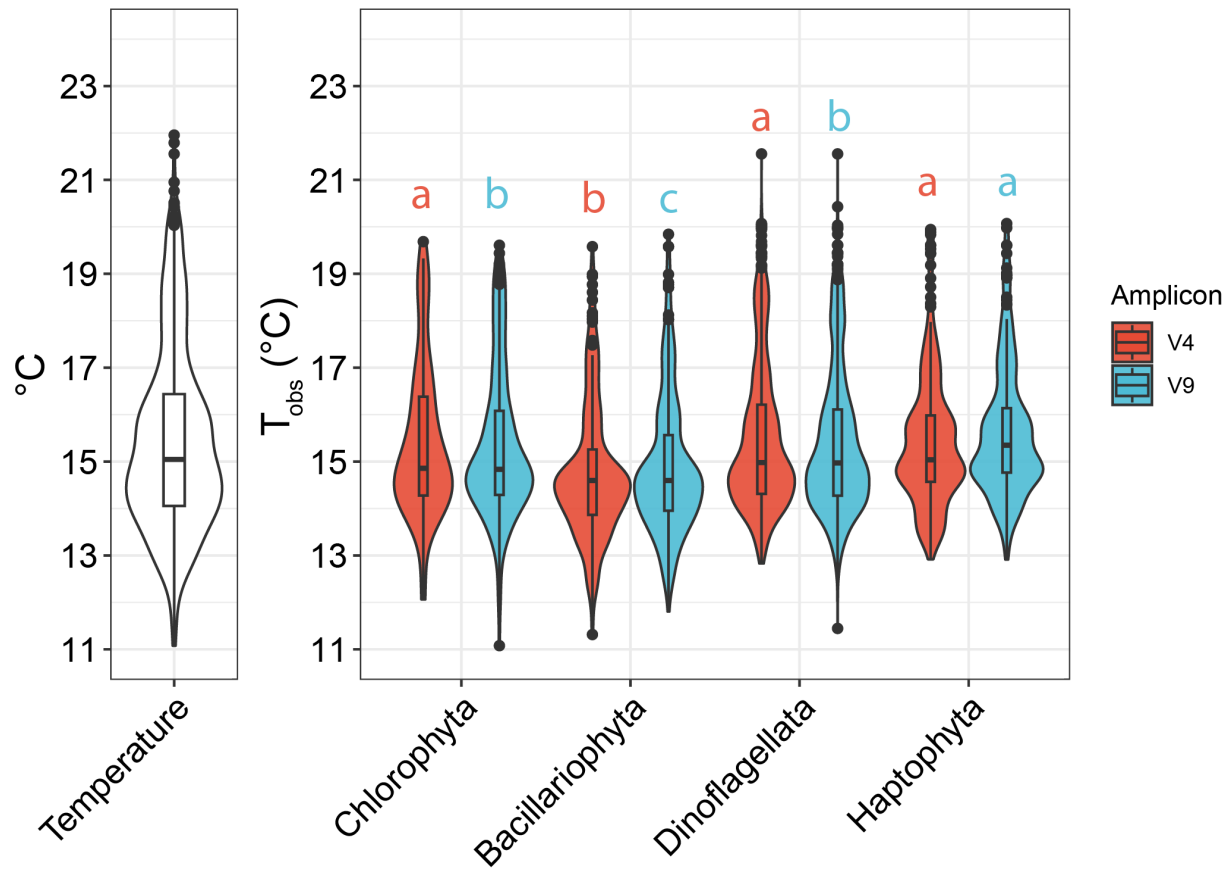
Supplemental Figure 4.5: Summed CLR-transformed diatom abundances versus environmental variables in both V4 (red) and V9 (blue) data with Spearman rank correlation coefficients shown above each panel. Correlations were computed with ALDEx2 using 1000 Monte Carlo instances and are all significant (adjusted $P < 0.05$).



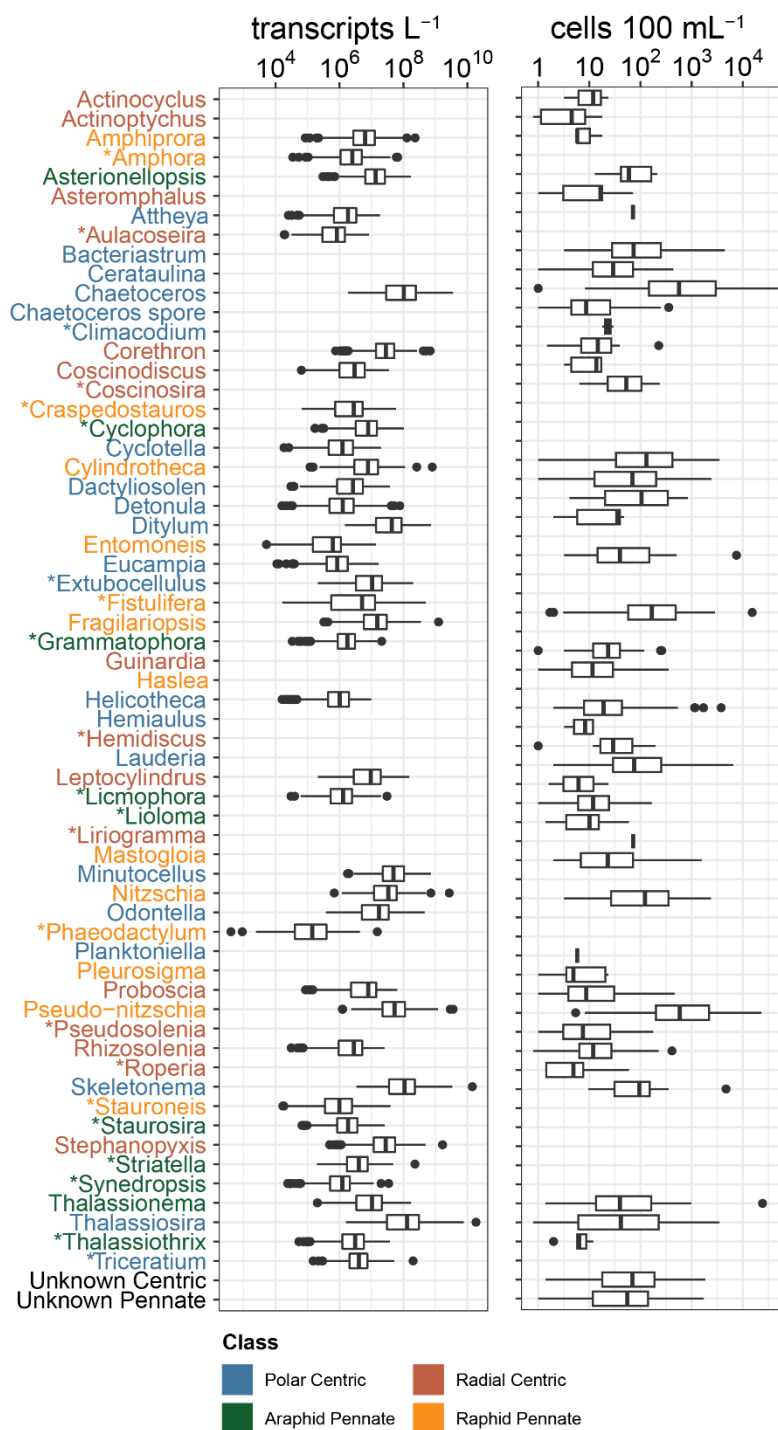
Supplemental Figure 4.6: Summed diatom absolute abundances (copies L^{-1}) versus environmental variables in both V4 (red) and V9 (blue) data with Spearman rank correlation coefficients. All correlations were significant (adjusted $P < 0.05$).



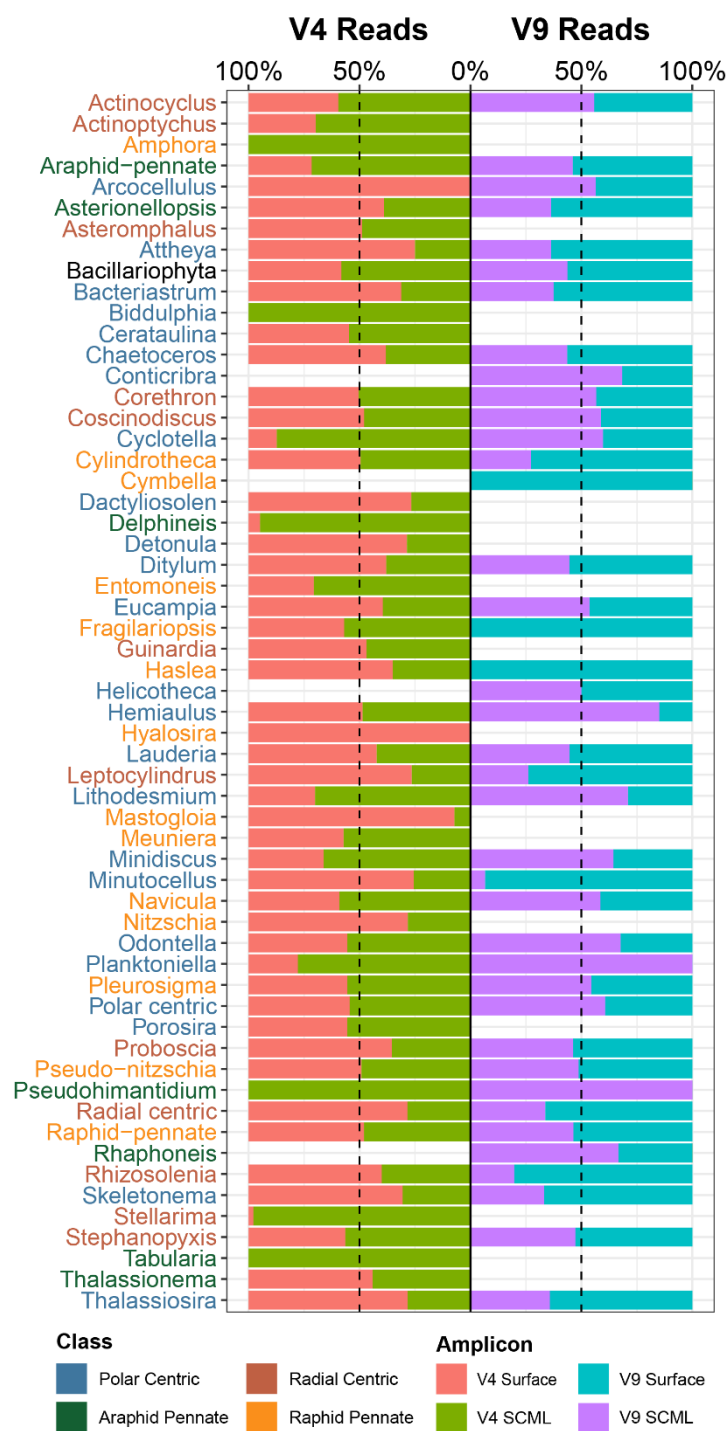
Supplemental Figure 4.7: (A) CLR-transformed relative abundances and (B) absolute abundances (copies L⁻¹) separated by season for Chlorophyta, Dinoflagellata, and Haptophyta with the V4 (red) and V9 (blue) data. Letters denote medians that are significantly different (adjusted *P*-value < 0.05, Kruskal-Wallis test). Black letters indicate that the significant differences are the same for V4 and V9 data. Otherwise, they are shown in red for V4 and blue for V9.



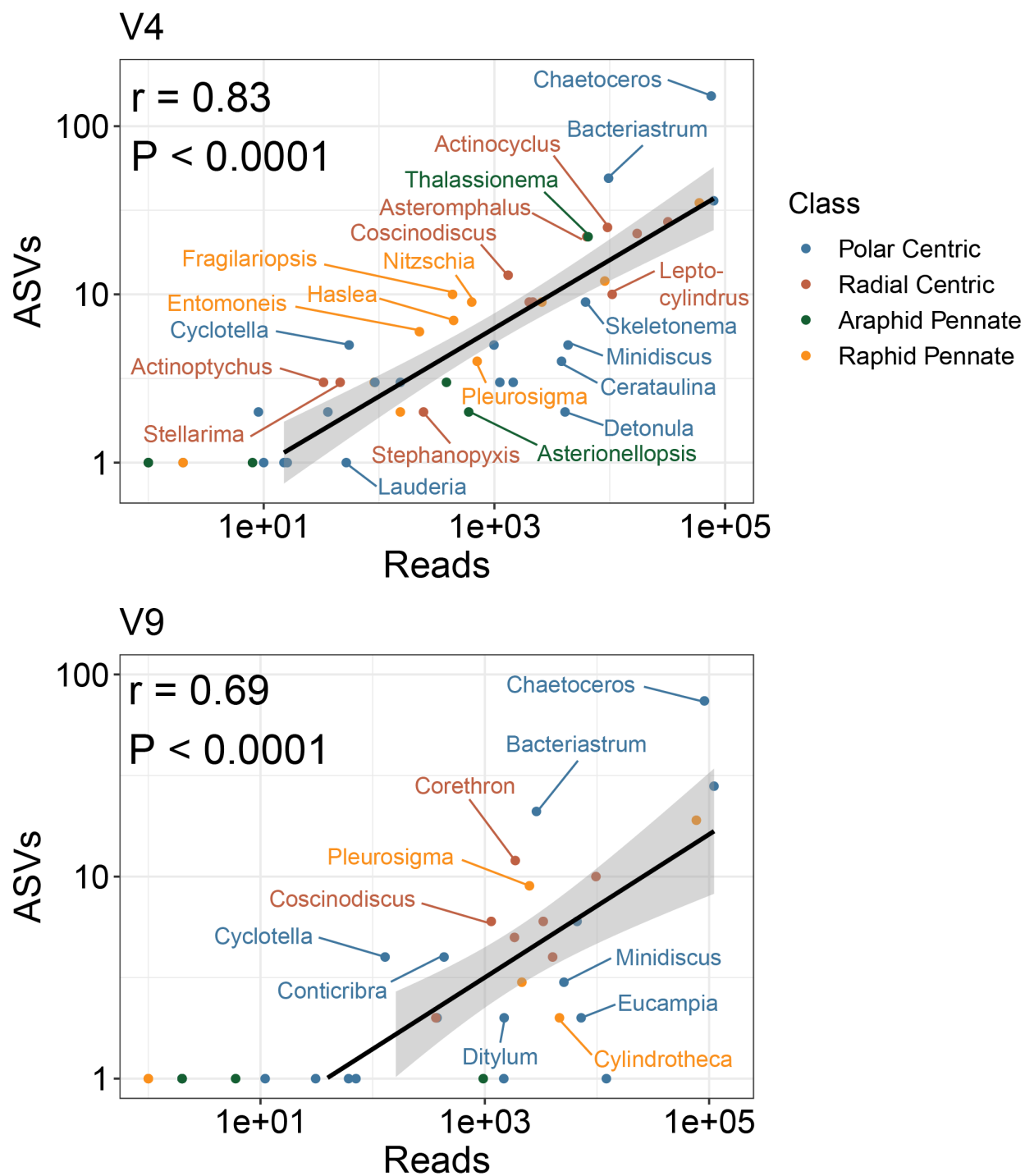
Supplemental Figure 4.8: Realized temperature niches (T_{obs}) for ASVs detected in 10 or more euphotic zone samples within dominant eukaryotic phytoplankton groups for the V4 (red) and V9 (blue) data. Letters denote medians that are significantly different (adjusted P -value < 0.05, Kruskal-Wallis test). The distribution of the underlying temperature measurements is also shown in the left panel.



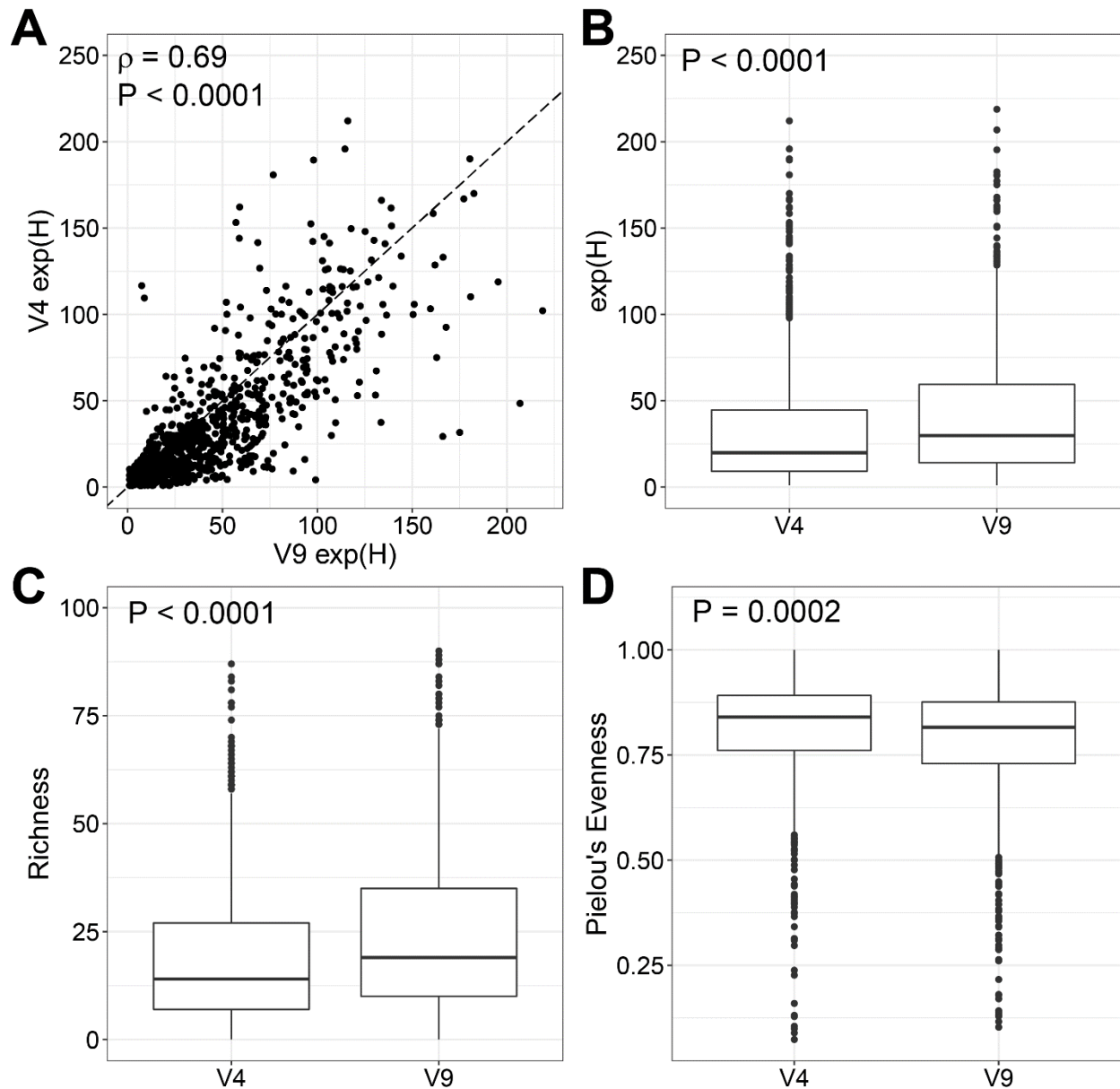
Supplemental Figure 4.9: Genus-level diatom abundances and richness (log scale) based on transcripts per L from the metatranscriptome (left) and microscopic cell counts (right, cells per 100 mL). Genera are color-coded by class as shown in the legend.



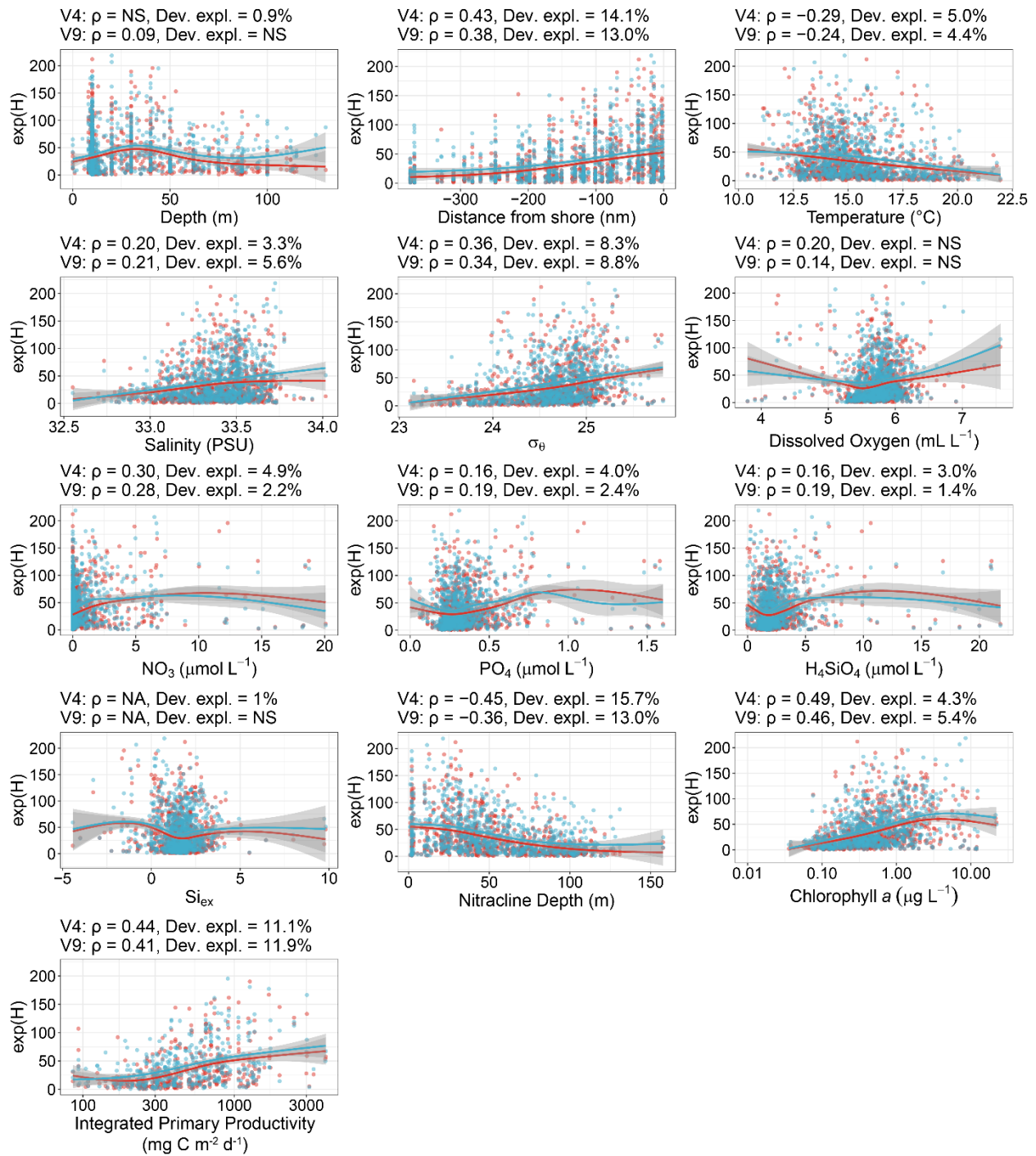
Supplemental Figure 4.10: The percentage of rarefied reads found in near-surface versus subsurface chlorophyll maxima layers (SCML) for each diatom genus and both the 18S-V4 (left) and 18S-V9 (right) data. Genera are color-coded by class levels as shown in the legend. ASVs without genus-level annotation are aggregated by the lowest taxonomic-level available.



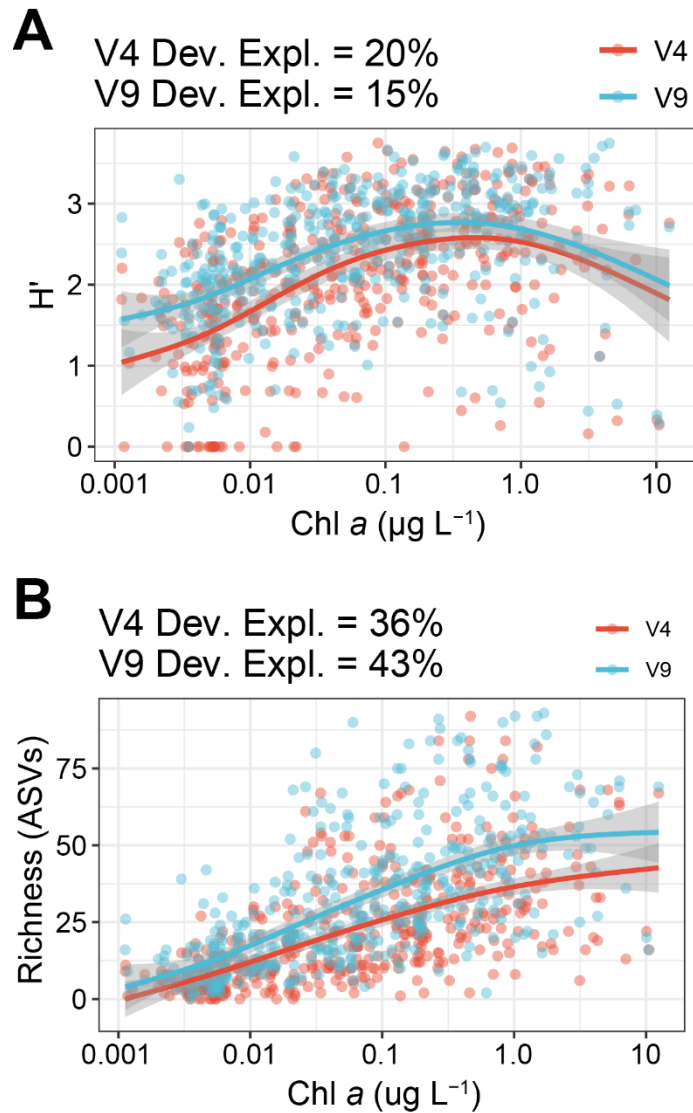
Supplemental Figure 4.11: Relationship between genus abundance and richness as shown in Fig. 4 for the V4 (top) and V9 (bottom) data. Points are colored by class as shown in the legend. The black line and gray shading represent the linear relationship and 95% confidence interval where total reads are greater than 10. Pearson correlation coefficients and P-values are shown in each panel.



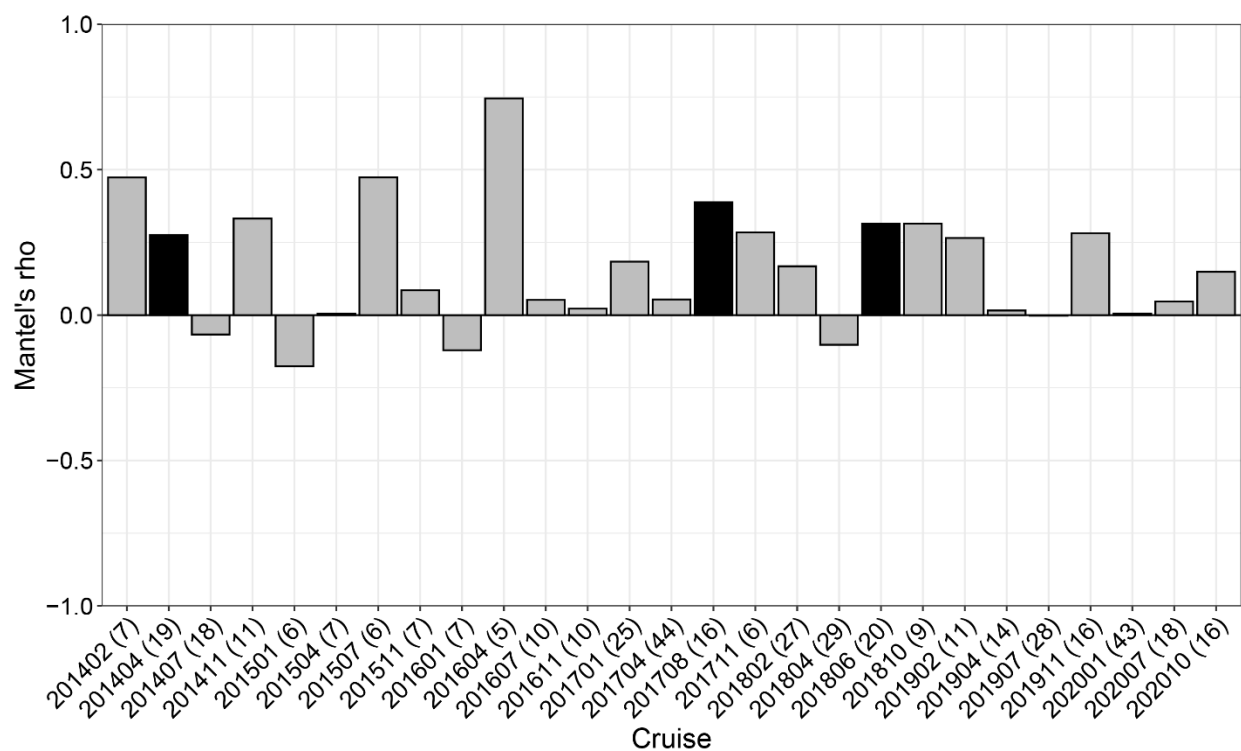
Supplemental Figure 4.12: Alpha diversity metrics between the diatom-only V4 and V9 data after rarefaction. The means within each metric are significantly different ($P < 0.001$, Student's t-test). (A) Pearson correlation between exponentiated Shannon index values. (B) Exponentiated Shannon index values shown as boxplots for each dataset. (C) Sample richness (number of diatom ASVs). (D) Pielou's evenness index.



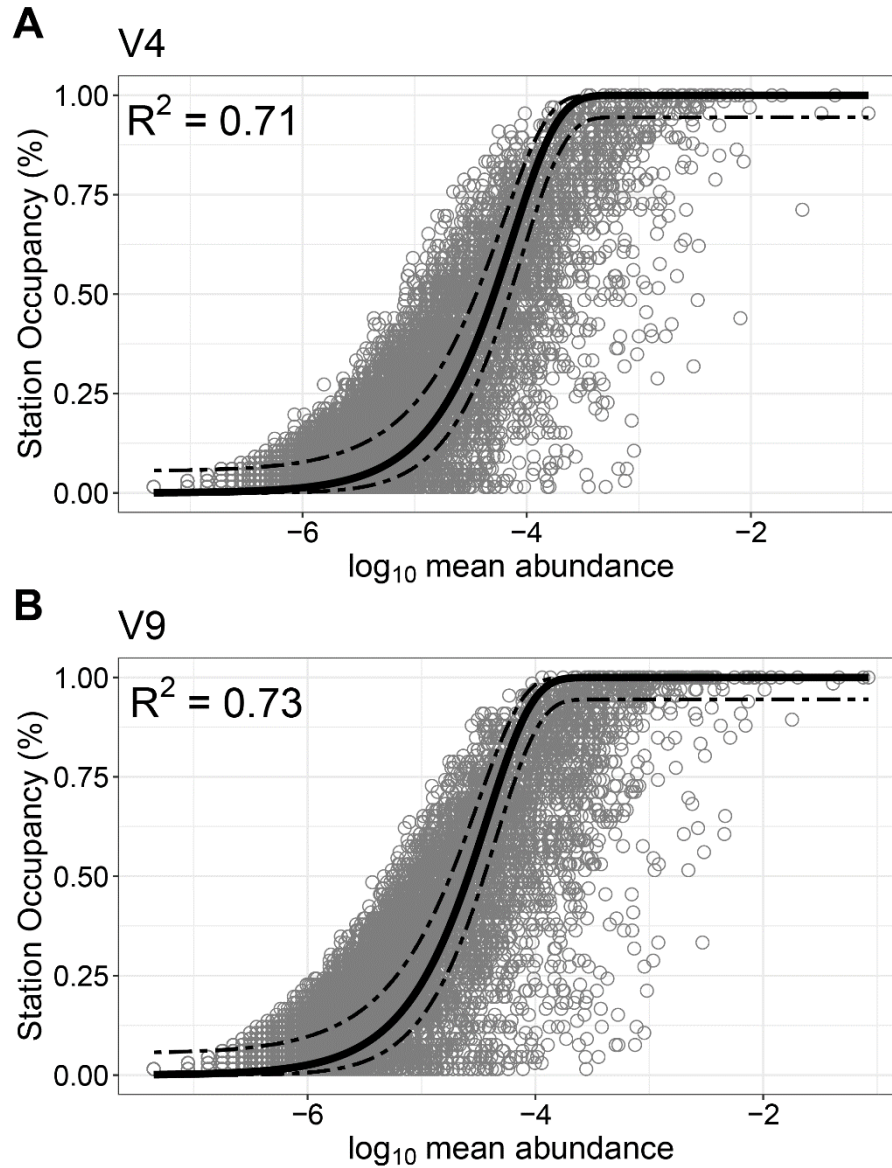
Supplemental Figure 4.13: Relationships between alpha diversity expressed as the exponentiated Shannon index and continuous variables for both the V4 (red) and V9 (blue) datasets. Generalized additive models (GAM) were used to generate trendlines with 95% confidence intervals. Spearman correlation coefficients (ρ) and the percent deviance explained by the GAM are shown above each plot. NS = not significant ($P > 0.05$).



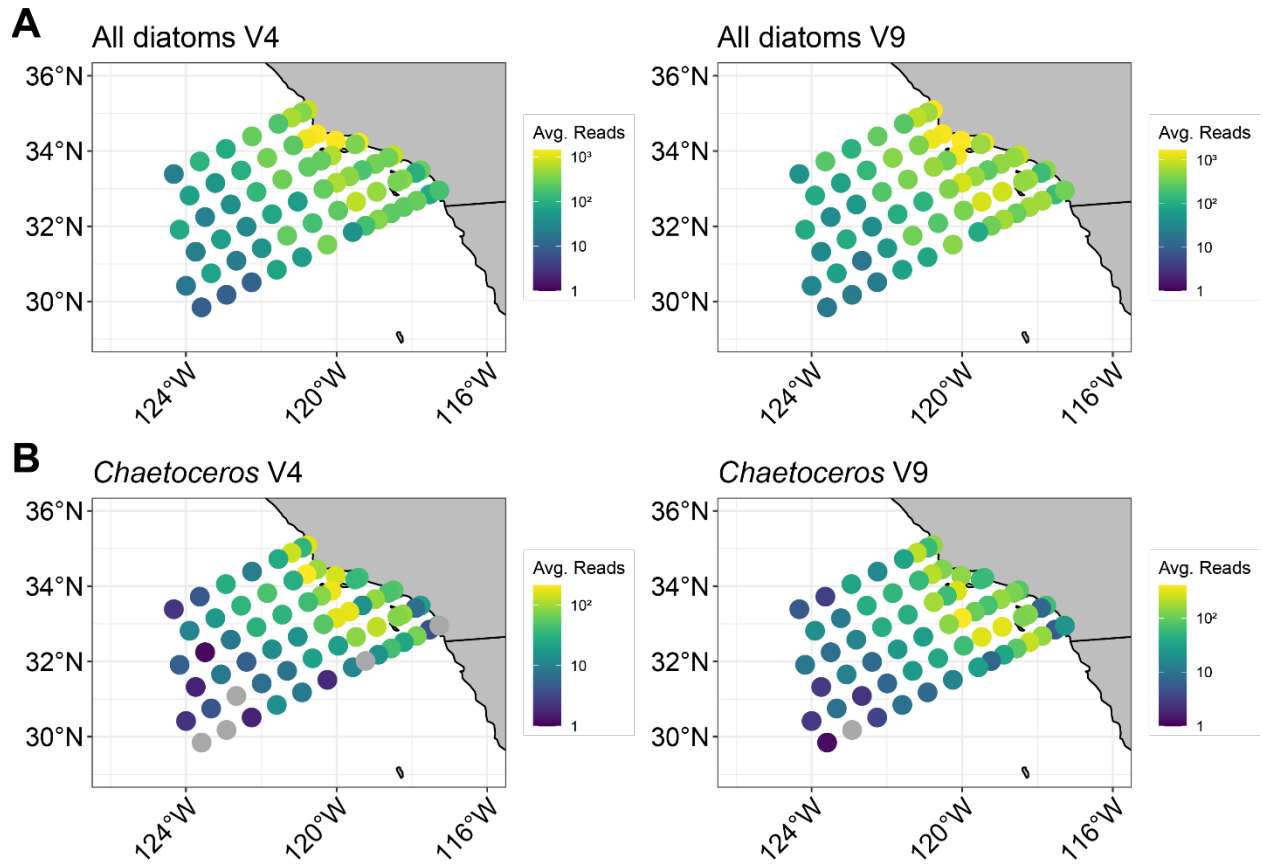
Supplemental Figure 4.14: Diatom diversity expressed as the (A) Shannon Index or (B) richness (number of ASVs) against taxon-specific biomass (μg chlorophyll a L^{-1}). Lines represent GAMs with 95% confidence intervals, and the deviance explained by each GAM is shown above each panel. Data are reproduced from Lampe et al. (2024).



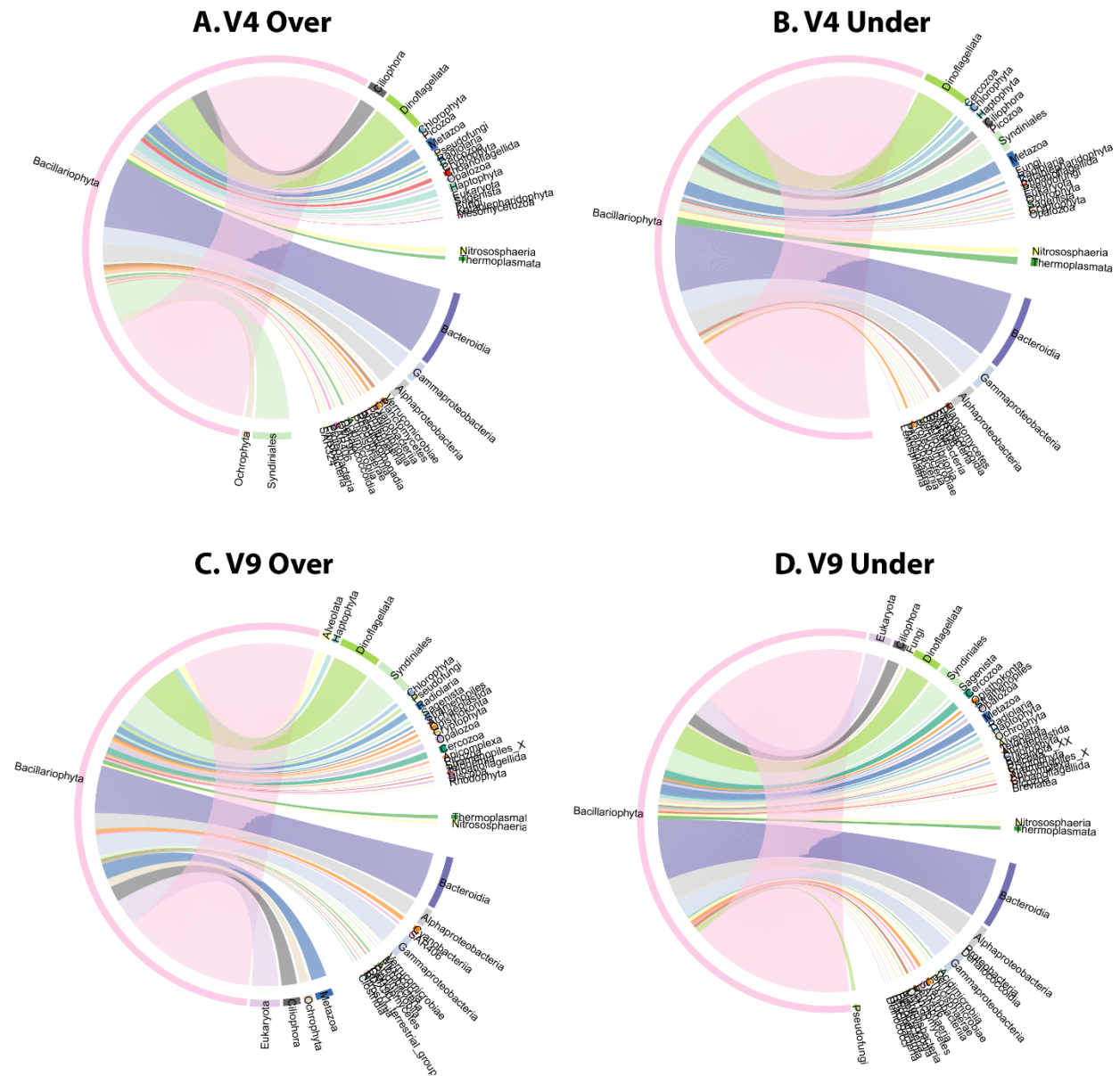
Supplemental Figure 4.15: Mantel test results between distance between stations and diatom community dissimilarity within each cruise. Cruises are labeled by ID (YYYYMM format) with the number of station ins parentheses on the x-axis. Significant correlations are shown in black (FDR < 0.05), otherwise correlations are shown in grey.



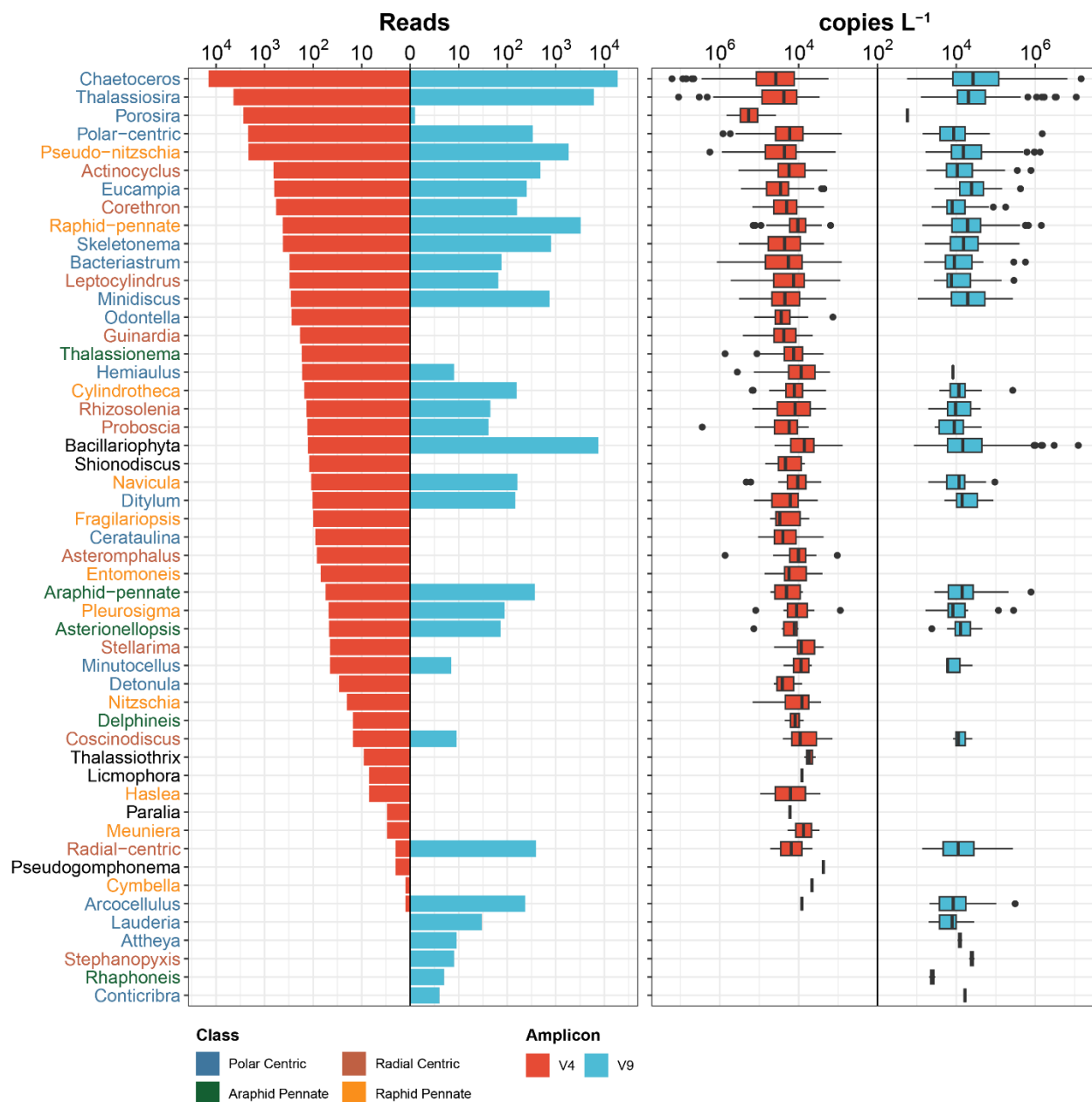
Supplemental Figure 4.16: Abundance-occupancy distributions for all (A) V4 and (B) V9 ASVs from the rarefied data. The solid line represents the fit of the neutral model, and the dashed line represented the 95% confidence interval around the model prediction.



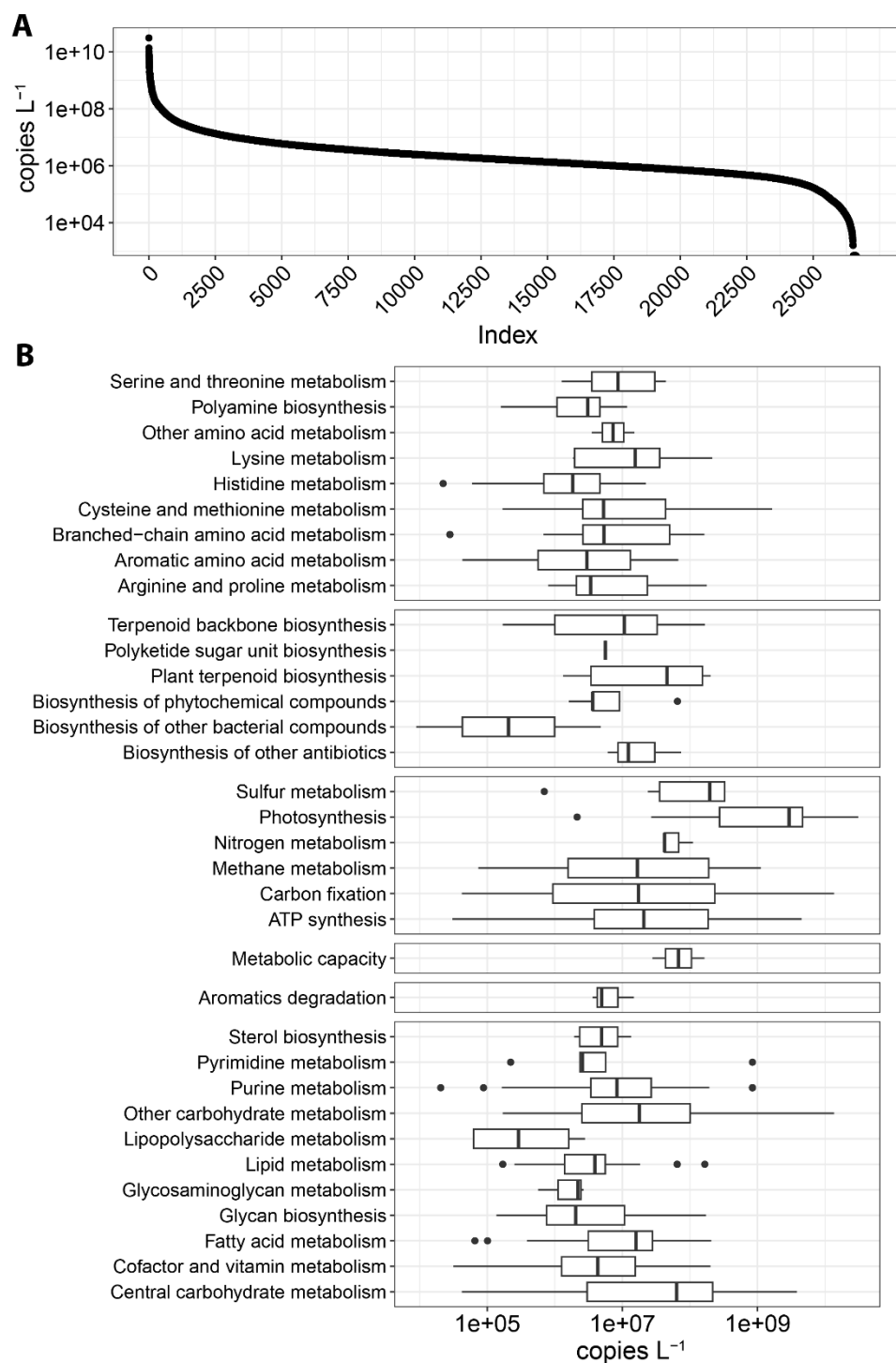
Supplemental Figure 4.17: The average read abundance on a log scale per station for (A) all diatom ASVs and (B) *Chaetoceros* ASVs that are predicted to be dispersal limited by the neutral model.



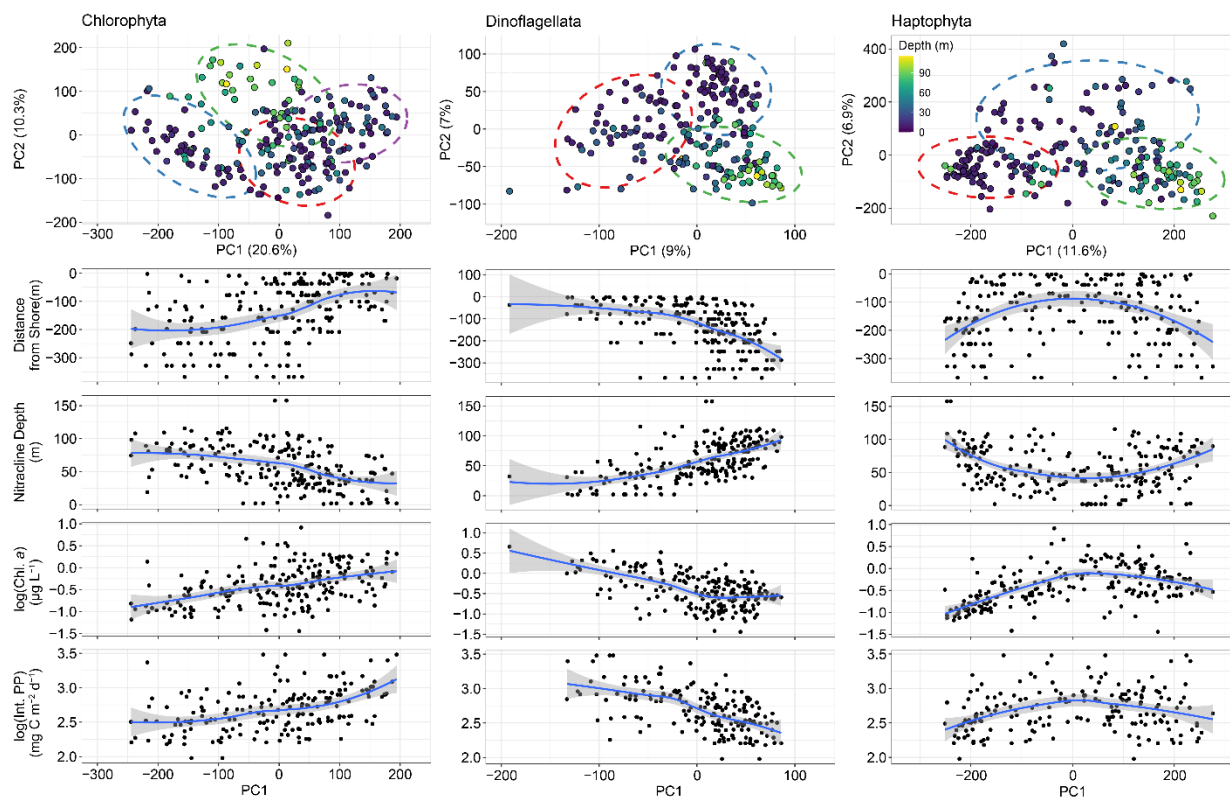
Supplemental Figure 4.18: Connections between (A, B) diatom V4 ASVs or (C, D) diatom V9 ASVs and other taxa as inferred from the co-occurrence network. Only diatom ASVs that are (A, C) above or (B, D) below the 95% confidence intervals of the neutral model are included.



Supplemental Figure 4.19: Genus-level diatom abundances within mesopelagic zone samples from (A) rarefied relative abundances and (B) absolute 18S abundances (copies L⁻¹). V4 data are shown in red and on the left, and V9 data are shown in blue on the right of each panel. ASVs without genus-level annotation are shown at the lowest taxonomic level possible.



Supplemental Figure 4.20: The absolute abundances of diatom gene clusters. (A) Gene clusters ordered by abundance. (B) Boxplots of diatom gene expression by KEGG module classes where applicable.



Supplemental Figure 4.21: Gene expression in other dominant phytoplankton groups. PCAs for each group with k-means clusters are shown above followed by distance from shore (m), nitracline depth (m), chlorophyll *a* on ($\mu\text{g L}^{-1}$) on a log scale, and integrated primary productivity ($\text{mg C m}^{-2} \text{ d}^{-1}$) on a log scale plotted against the first principal component. Colors in the PCA represent depth (m), and lines represent generalized additive models with associated 95% confidence intervals.