Relationships between phytoplankton pigments and DNA- or RNA-based abundances support ecological applications

Supplementary Information

Robert H. Lampe[1,2], Ariel J. Rabines[1,2], Steffaney M. Wood[1,2,3], Anne Schulberg[1,2], Ralf Goericke[1], Pratap Venepally[2], Hong Zheng[2], Michael R. Stukel[4], Michael R. Landry[1], Andrew D. Barton[1,5], Andrew E. Allen[1,2]

[1]Integrative Oceanography Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, 92093 USA
[2]Microbial and Environmental Genomics, J. Craig Venter Institute, La Jolla, CA, 92037, USA
[3]Center for Marine Biotechnology and Biomedicine, Scripps Institute of Oceanography, University of California San Diego, La Jolla, CA, 92093 USA
[4]Earth, Ocean, and Atmospheric Science Department, Florida State University, Tallahassee, FL, 32304 USA
[5]Department of Ecology, Behavior and Evolution, University of California San Diego, La Jolla, CA, 92093 USA

Correspondence to: Andrew E. Allen (aallen@ucsd.edu)

Figure S1. (A) Temperature (°C) and (B) nitrate concentrations (µmol L^{-1}) among seasons for the RNA samples.

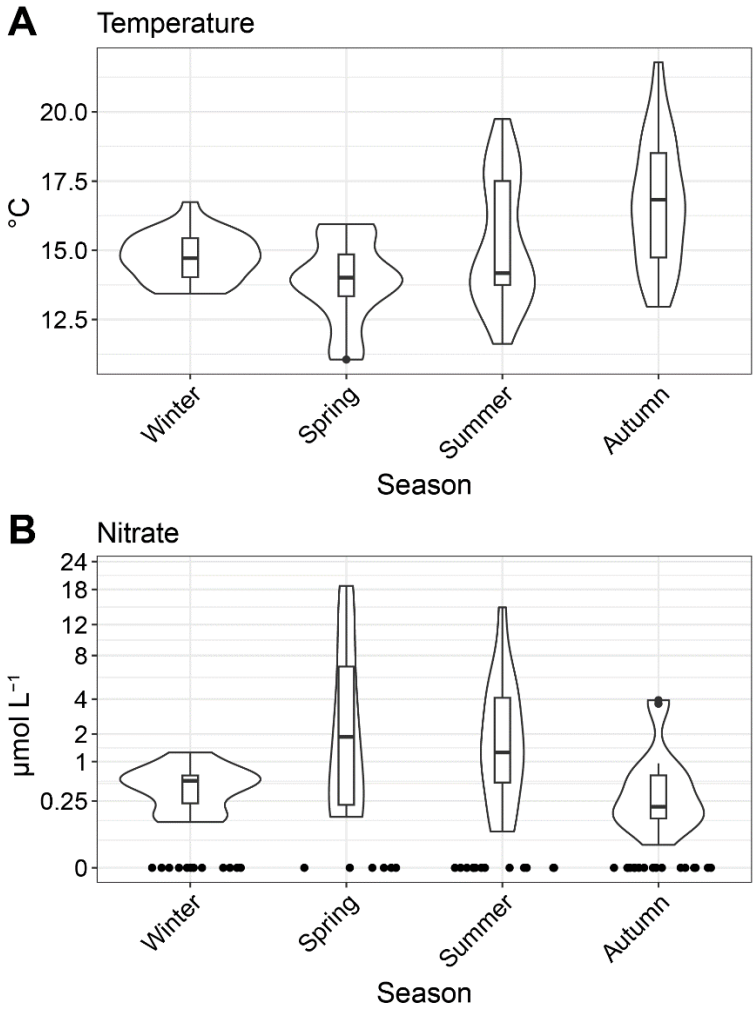

Figure S2. Cell abundances (cells mL^{-1}) for *Prochlorococcus* (Pro) and *Synechococcus* (Syn) as (A) distributions and (B) sample-by-sample comparisons where the dashed line represents a one-to-one relationship.

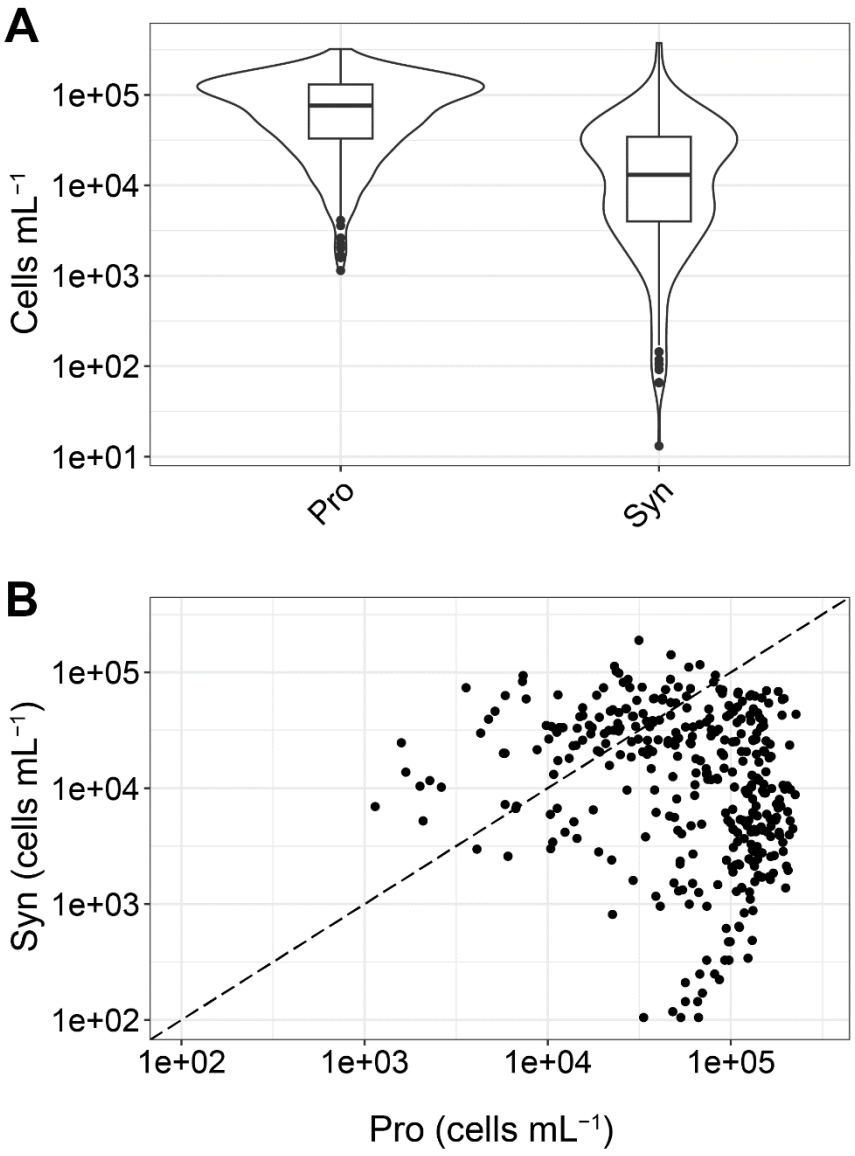

Figure S3. Pearson correlations among absolute abundances of phytoplankton groups and diagnostic pigments for (A) 18S-V4 and (B) 18S-V9 data. Correlations are not shown where $P > 0.05$.

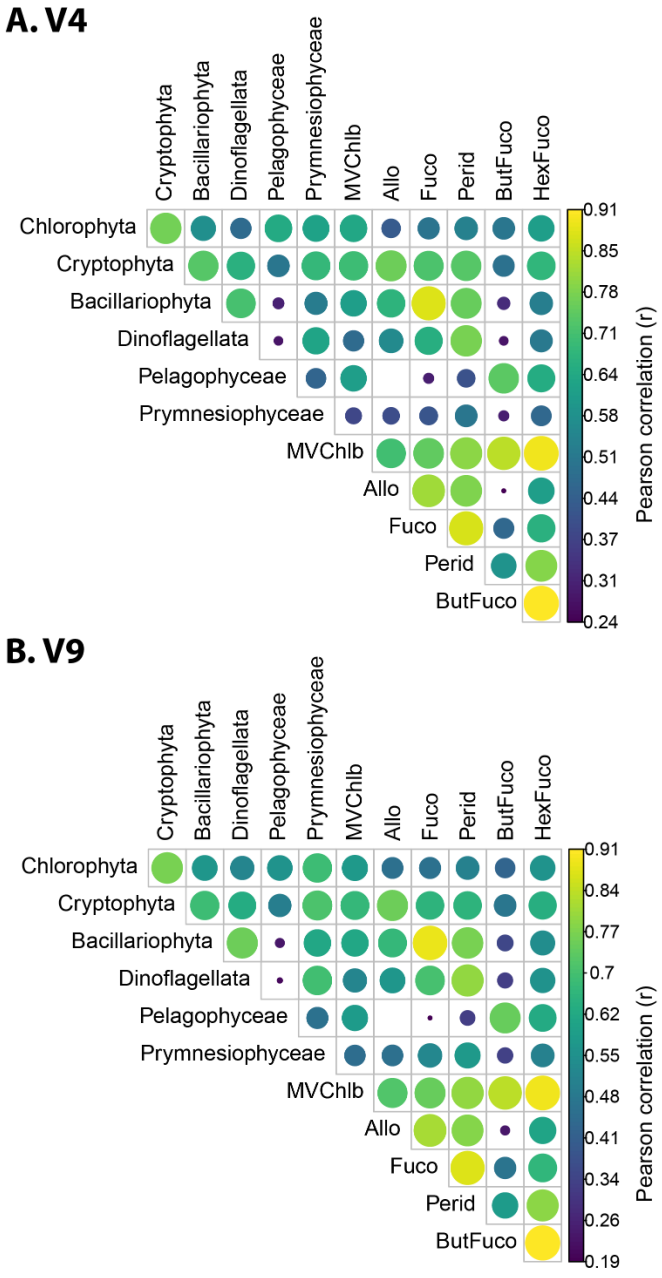

Figure S4. Correlations between 18S-V4 (red) or 18S-V9 (blue) absolute abundances (copies L^{-1}) and taxon-specific contributions to total chlorophyll estimated with phytoclass (µg L^{-1}) for each major eukaryotic phytoplankton group. Pearson correlations are significant ($P < 0.05$) and coefficients are displayed in each panel. Lines with 95% confidence intervals shows linear models with significant relationships between variables ($P < 0.05$).

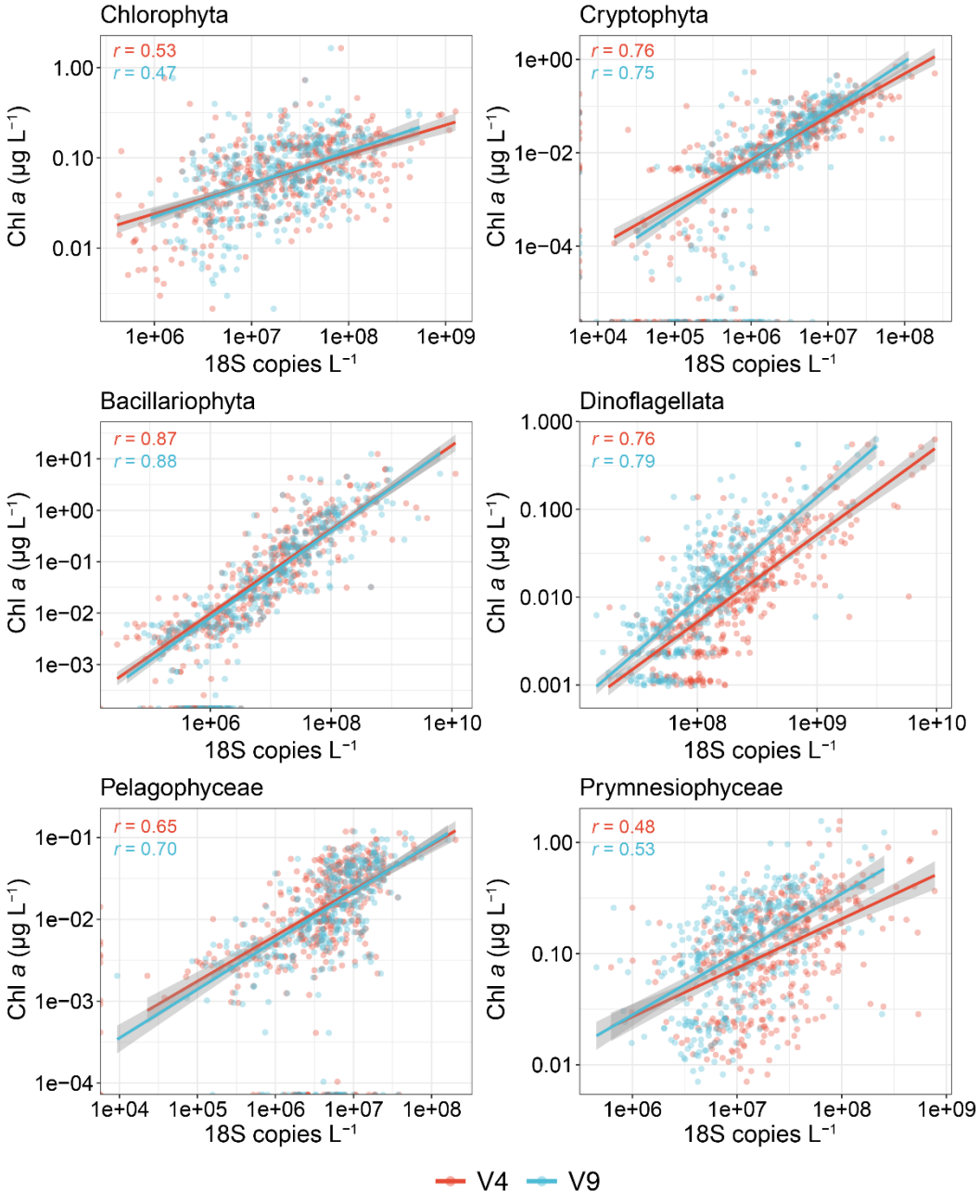

Figure S5. Linear regressions for depth (m) in predicting residuals from linear models on the absolute abundance and diagnostic pigments shown in Fig. 2B. Coefficients of determination (R^2) and regression lines are shown when $P < 0.05$.

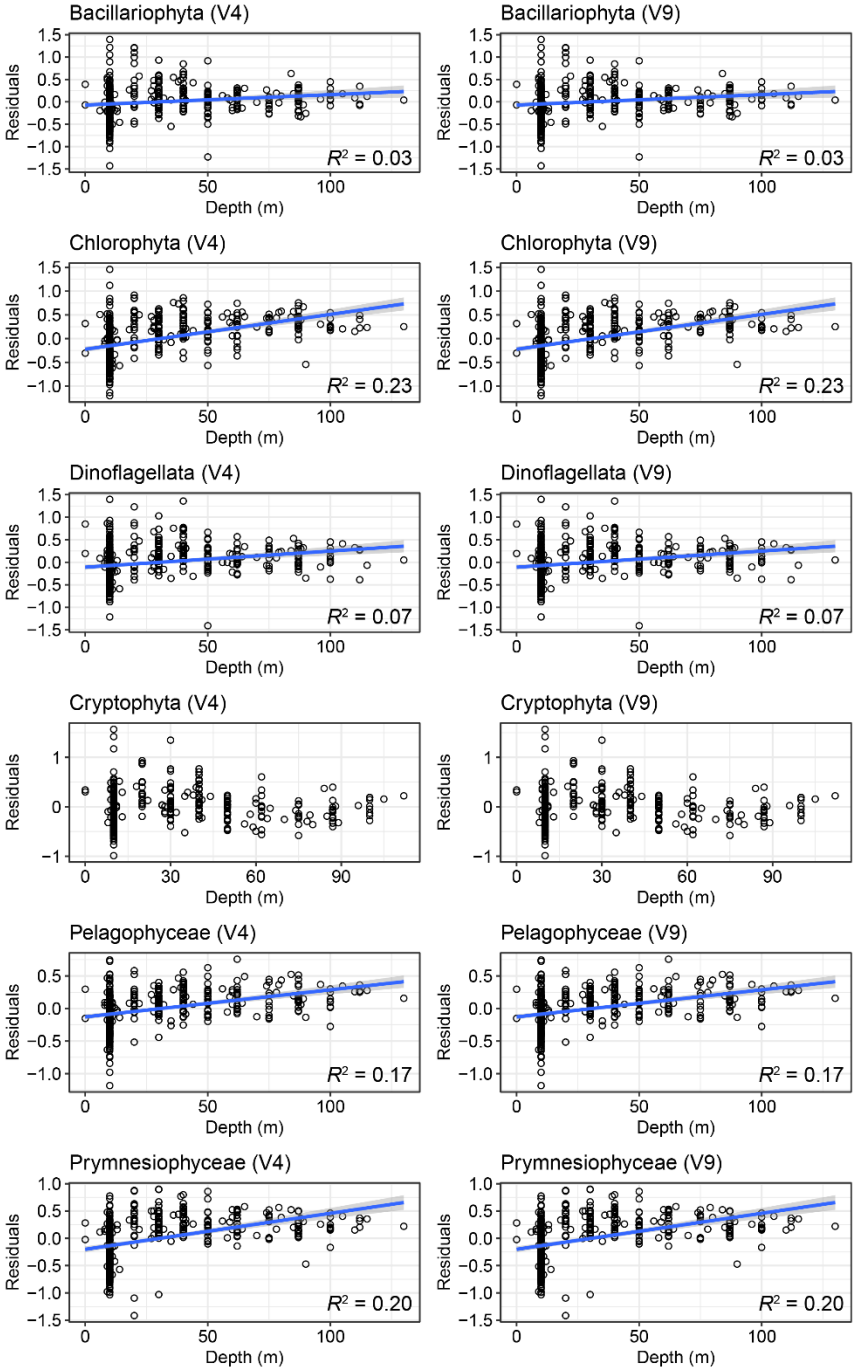

Figure S6. Relative abundances based on total copies L^{-1} of Dinoflagellata ASVs on the (A) order level including Syndiniales and (B) genus level when Syndiniales are removed.

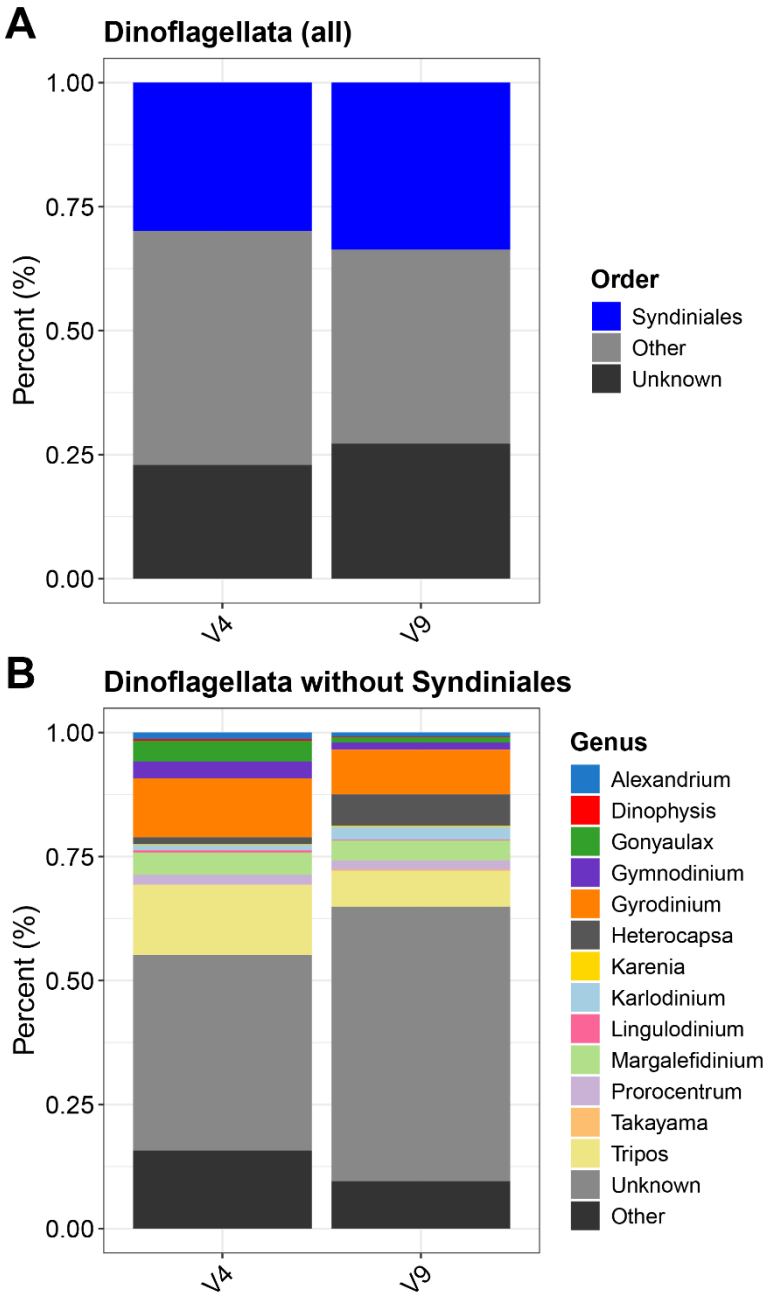

Figure S7. Correlations between diagnostic pigment concentrations and the absolute abundances of mRNA (transcripts L^{-1}) for (A) all cyanobacteria and zeaxanthin and (B) *Prochlorococcus* and divinyl chlorophyll *a* (DVChla). Only samples from the upper 20 m are included. Pearson correlations are significant ($P < 0.05$) and coefficients are displayed in each panel. Lines with 95% confidence intervals show linear models with significant relationships between variables ($P < 0.05$).

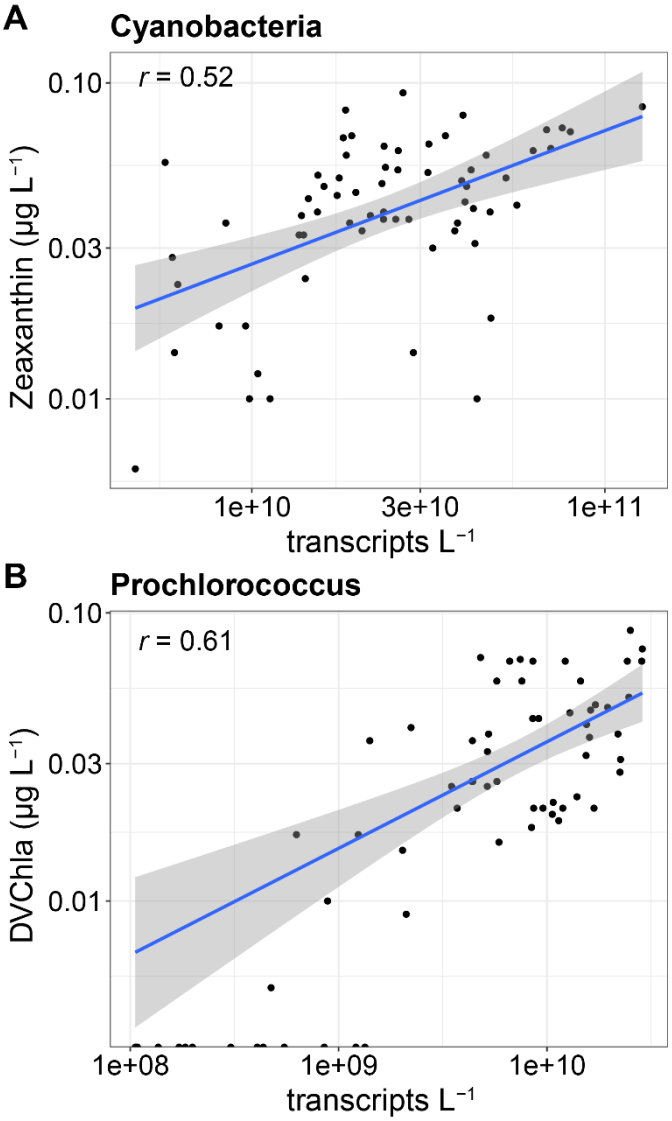

Figure S8. Productivity-diversity relationships in the region with net primary productivity (NPP) and both (a) the Shannon Index and (b) richness (ASVs) using both the 18S-V4 (red) and 18S-V9 (blue) data. Lines represent GAMs and corresponding 95% confidence intervals where significant ($P < 0.05$). 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.

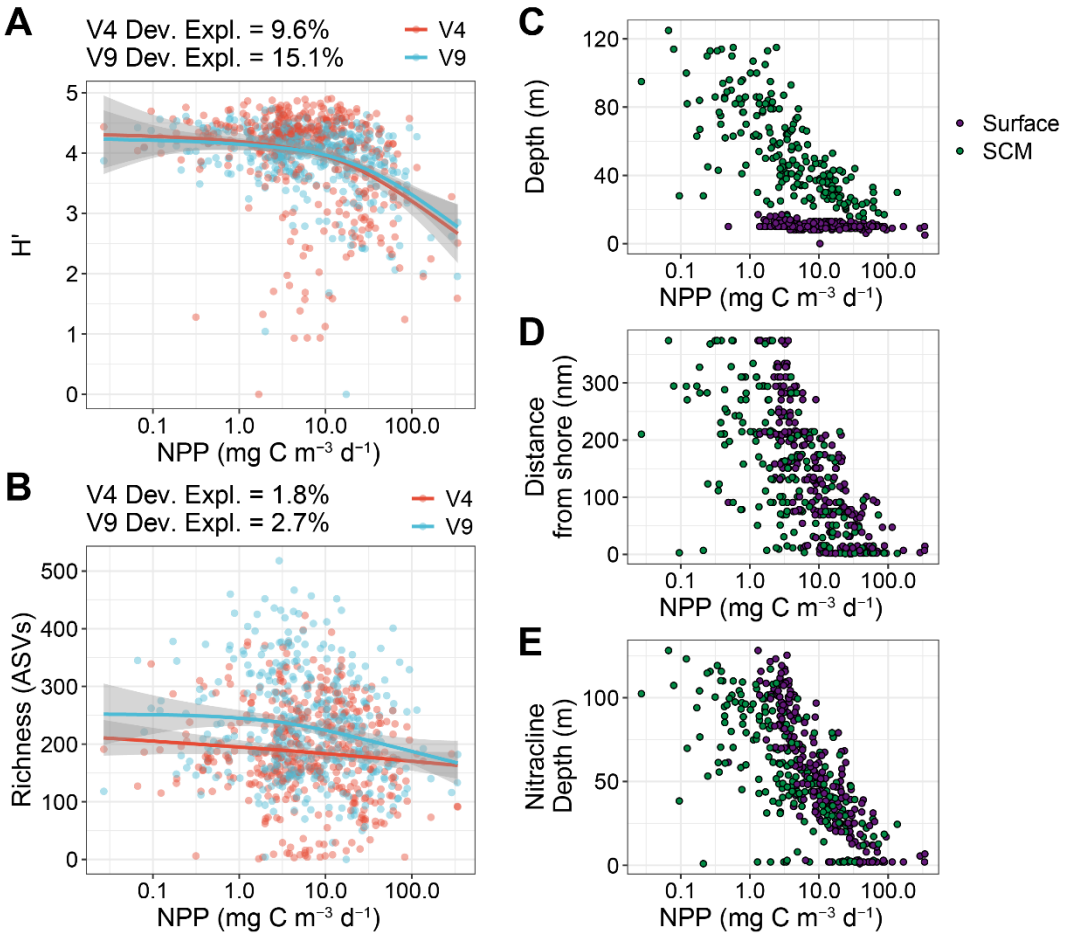

Figure S9. Productivity-diversity relationship in the region for (A) all eukaryotic phytoplankton groups or (B) all eukaryotic groups and cyanobacteria with biomass (chlorophyll a concentrations) as a proxy for productivity and diversity expressed as the richness (number of ASVs) for both the 18S-V4 (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.

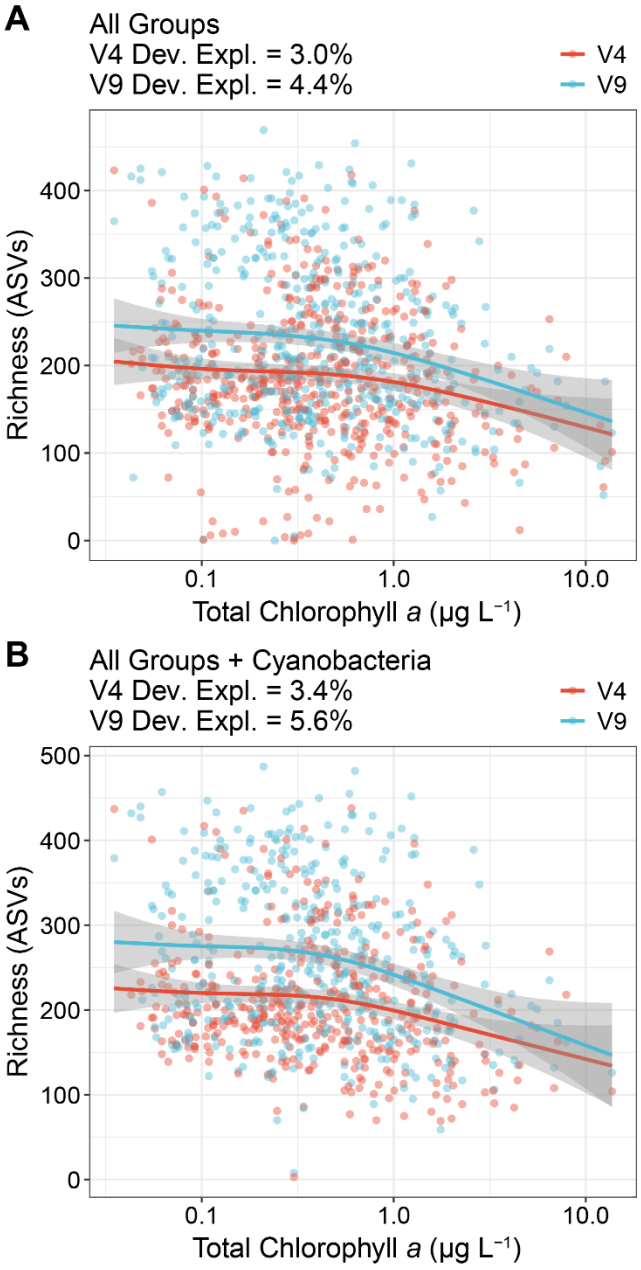

Figure S10. Productivity-diversity relationships for each individual phytoplankton groups using richness (ASVs) against taxon-specific biomass estimated with phytoclass. Lines represent GAMs and corresponding 95% confidence intervals where significant ($P < 0.05$). The deviance explained by each GAM or "NS" for not significant ($P > 0.05$) is shown above each panel for each amplicon.

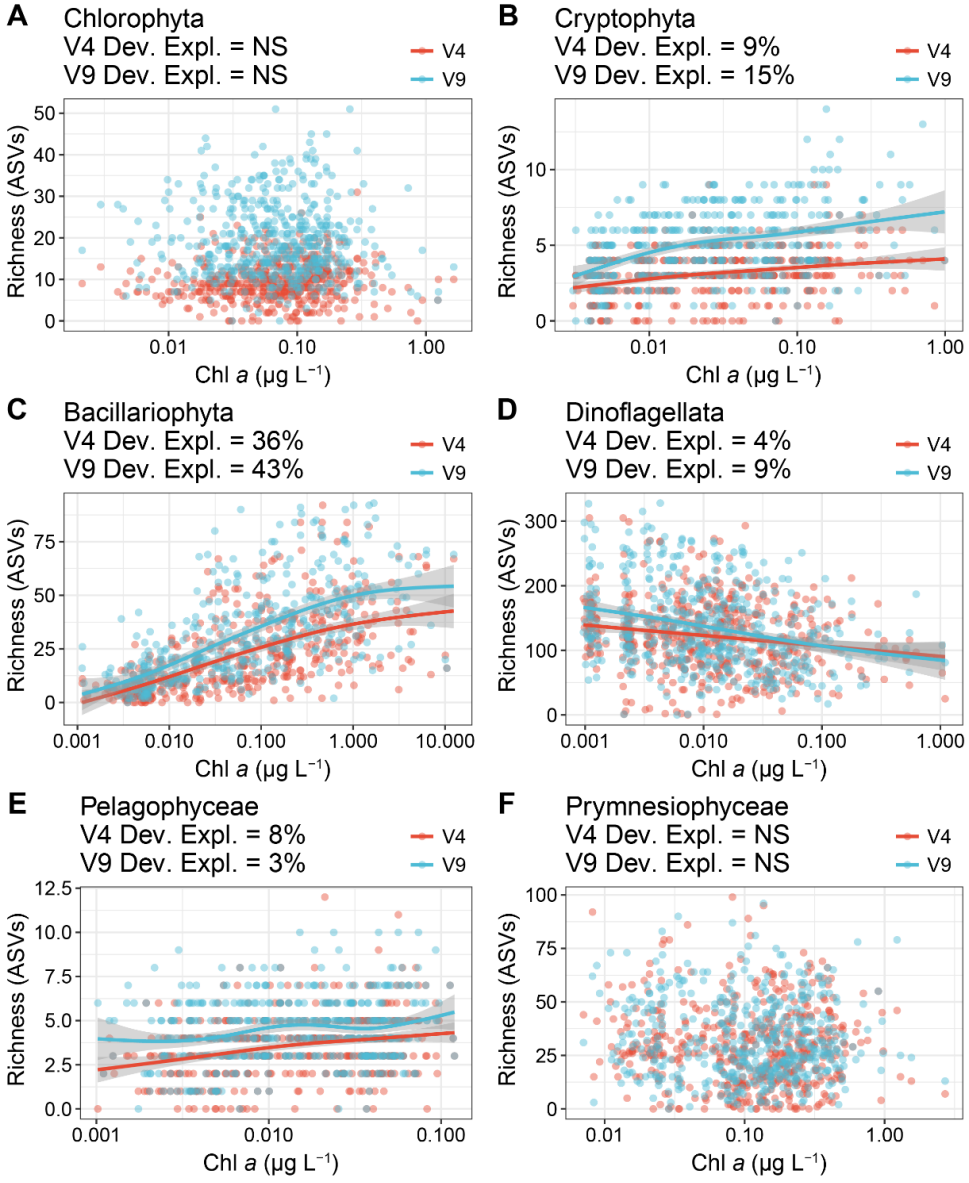

Figure S11. Productivity-diversity relationships in the region for all eukaryotic phytoplankton groups except for dinoflagellates with biomass (chlorophyll a concentrations) as a proxy for productivity and diversity expressed as the Shannon Index (H') for both the 18S-V4 (red) and 18S-V9 (blue) data. 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.

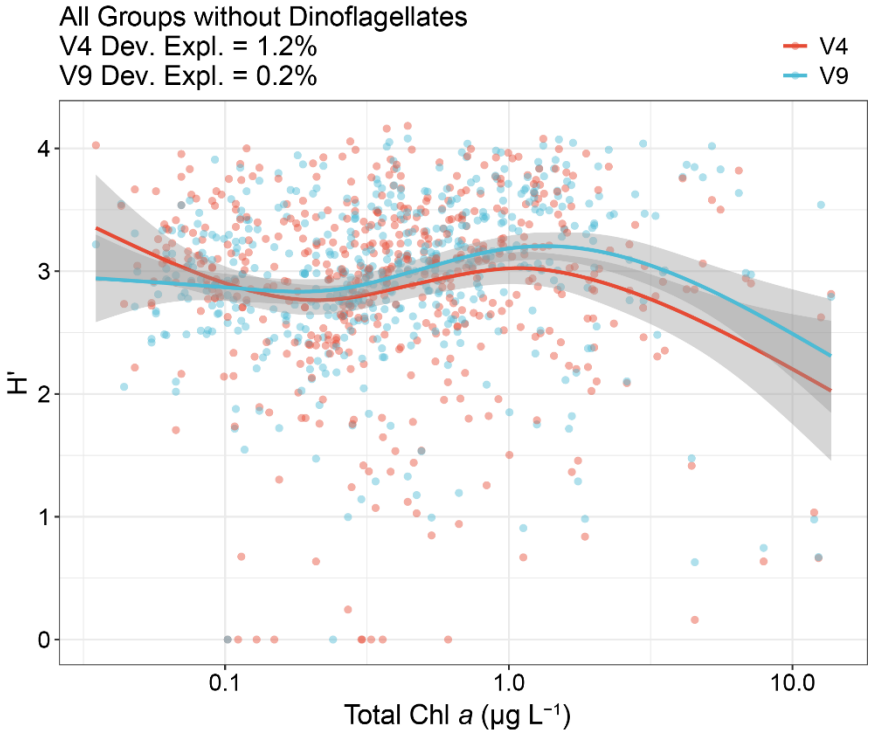

Figure S12. Diatom-to-dinoflagellate ratio for the (A) 18S-V4 and (B) 18S-V9 data against total chlorophyll *a* concentrations. The line and 95% confidence interval represent the GAM ($P < 0.0001$) that explains 15.0% (V4) or 18.2% (V9) of the deviance.

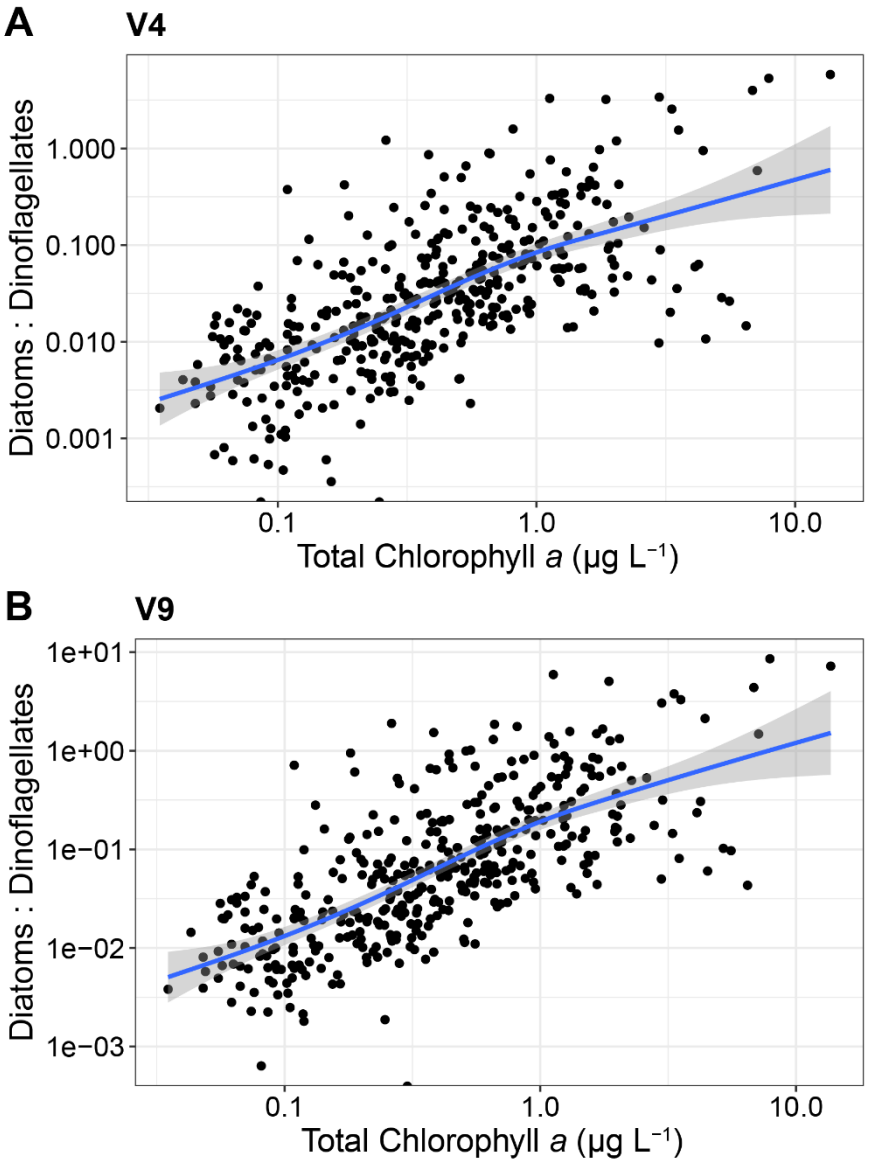

Figure S13. Dinoflagellate genus-level abundances against (A) total dinoflagellate copies L^{-1}, (B) total chlorophyll *a* concentrations, and (C) peridinin concentrations.

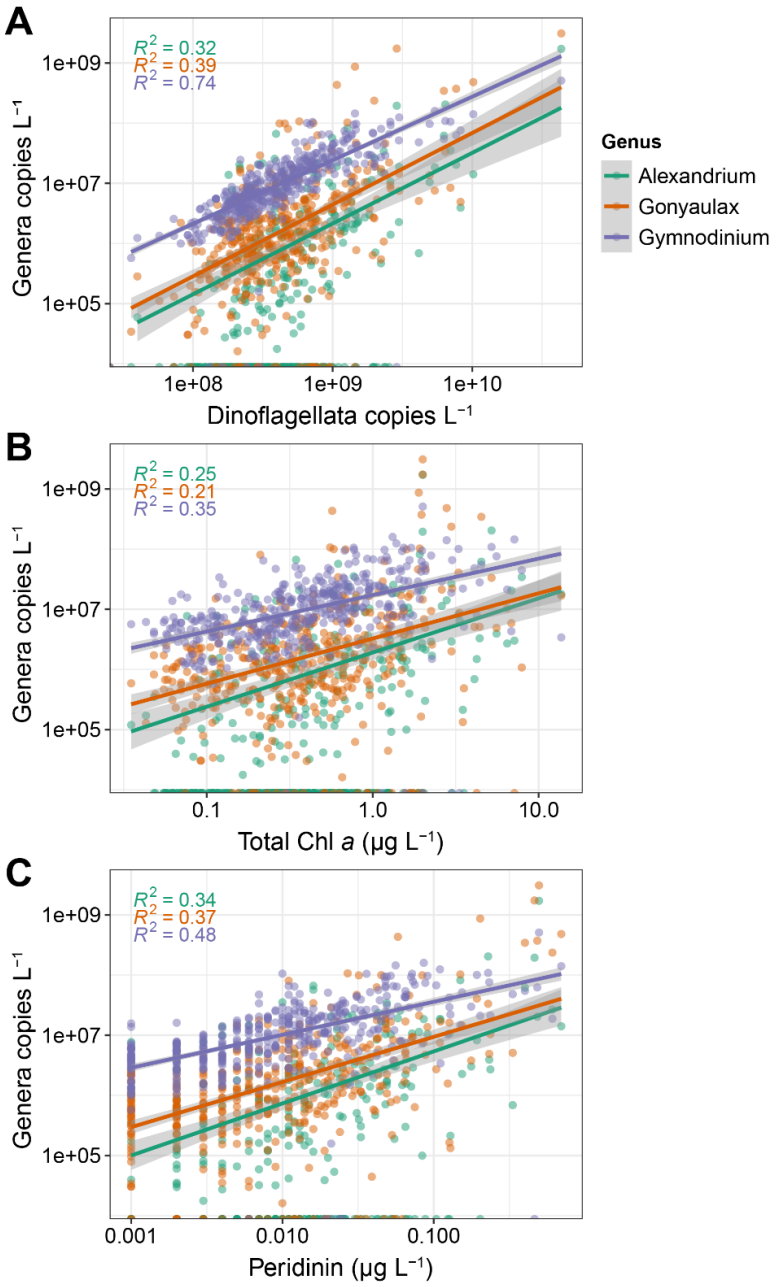

Table S1. Pearson correlation coefficients for dinoflagellates with different groups removed and peridinin.

	Amplicon			
	Relative		Absolute	
Removed Group	**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