Supplementary Information

Relating sinking and suspended microbial communities in the California Current Ecosystem: Digestion resistance and the contributions of phytoplankton taxa to export

Bellineth Valencia,[1] Michael R Stukel,[2] Andrew E Allen,[1,3] John P McCrow,[3] Ariel Rabines,[3] Brian Palenik[1] and Michael R Landry[1]

[1]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA; [2]Earth, Ocean, and Atmospheric Science Department, Florida State University, Tallahassee, FL, USA and [3]Microbial and Environmental Genomics, J Craig Venter Institute, La Jolla, CA, USA

Experimental Procedures

Cruise plan

Samples and environmental data were collected at four sites in the southern California Current Ecosystem (CCE) on the *R/V Sikuliaq* as part of the California Current Ecosystem – Long-Term Ecological Research (CCE-LTER) process cruise P1604 (19 April – 12 May 2016). Collections were done at the end of anomalously warm conditions (sea surface temperature anomalies >2°C) that prevailed in the region from late 2013 to 2016. Initially, the anomalously warm water, known as the "blob", extended southward from Alaska and affected the North Pacific (Bond *et al.,* 2015). In the subsequent 2015-2016 El Niño event, the warm-water conditions extended northward from the equatorial region (McClatchie *et al*., 2016). In the CCE, productivity tends to decrease during El Niño events, and plankton communities shift toward a higher dominance of smaller subtropical and open-ocean taxa (Chavez *et al.,* 2002; Peterson *et al.,* 2002; Peterson *et al.,* 2017). The aim of the cruise was to evaluate the effects of the 2015–2016 El Niño event on the pelagic ecosystem.

At each site, we conducted experimental cycles following water parcels over 3-5 days using a surface drifter with holey sock drogue at 15 m. Sampling began at midnight with the deployment of a sediment trap and an additional experimental drifter array (Ohman *et al.,* 2013). As we followed the arrays, the euphotic zone (0.1% surface irradiance) was characterized by taking profiles twice a day (02:00 and 12:00) with a Conductivity Temperature Depth (CTD) in a Niskin-bottle rosette, collecting discrete water samples at 6–8 depths for chlorophyll (Chl*a*), nutrients, primary production (PP), and metabarcoding. Sampling ended with the recovery of the experimental (~ 02:00) and sediment trap arrays (~ 06:00). Due to bad weather, sampling at the offshore site (OO) was interrupted, and the sediment trap array drifted for a longer period of time (Table S1).

Sediment traps

Sinking particulate matter was collected at each site in VERTEX-style sediment traps that were attached below the drogued drifter array (Table S1). Trap arrays consisted of 12 particle interceptor traps deployed at each of two depths, the base of the euphotic zone and 150 m. Each particle interceptor trap had an inner diameter of 70 mm, an aspect ratio of 8:1 (height:diameter), and baffle tubes on top to minimize resuspension during recovery

2

(Knauer *et al.,* 1979; Stukel *et al.,* 2013a). Here, we focus the analyses on the sinking particles collected in the deep sediment traps at 150 m from where two particle interceptor traps were assigned for metabarcoding analyses of the microbial communities. The tubes were filled before deployment with either 2.2 liters of a brine or a RNA later solution (Table S1). The brine solution consisted of 0.1 µm filtered seawater and 50 g l^{-1} of NaCl, creating a density interface to prevent mixing with in situ water (Stukel *et al.*, 2013a). The RNA later, made following the protocol described by Fontanez *et al.* (2015), was used to reduce DNA degradation in the traps. Briefly, 40 ml 0.5M of EDTA, 25 ml 1M sodium citrate, and 700 grams of ammonium sulfate were combined with ultrapure water. The solution was heated and stirred until the ammonium sulfate was dissolved. The pH of the solution was adjusted to 5.2 using sulfuric acid and particles were removed from the solution by filtering through a Sterivex filter (0.2 µm) with a peristaltic pump. We filled the entire particle interceptor traps with the RNA later solution.

After recovering the trap array, sample processing followed the protocol described by Stukel *et al.* (2013a): the depth of the salinity interface was established, the overlying water was removed by gentle suction, and the water was filtered through a 47-mm diameter Nitex screen (200-µm pore size) to remove mesozooplankton swimmers that were carefully checked under a dissecting microscope (20X magnification). Non-swimmer particles larger than 200-µm were kept in the Nitex screen, placed in a 2-ml screw-cap cryogenic vial, flash-frozen in liquid nitrogen, and stored at -80°C. The remaining water and particles were filtered through Sterivex filters (0.2 µm) with a peristaltic pump, flash frozen, and stored at -80°C. The mean volume of water filtered was 1.7 liters (range: 1.37 – 2.10 liters) for tubes filled with the brine solution and 1.8 liters (range: 1.32 – 2.10 liters) for tubes filled with the RNA later solution.

Water column
Samples for metabarcoding were collected from the mixed layer, the lower euphotic zone and at the deep trap depth (150 m) to evaluate which microbes from the water column contributed to sinking particles exported from the euphotic zone (Table S1). 280 ml (200-µm Nitex screen) or 650 ml of seawater (500-µm Nitex screen) were pre-screened to remove mesozooplankton prior to filtration through a 25-mm diameter 0.2-µm Supor

membrane filters (Pall Corporation). Once the water was filtered, the filters were folded in half, placed in 2-ml screw-cap cryogenic vials, flash-frozen in liquid nitrogen, and stored at -80°C until analysis.

Library construction and sequencing

Environmental DNA from the water column and sediment trap samples was extracted using the NucleoMag 96 Plant kit (Macherey Nagel) following the manufacturer's instructions. Supor membrane filters (water column samples) and Nitex screen filters (sediment trap particles > 200 µm) were placed directly in the lysis buffer. In the case of the Sterivex filters (sediment trap samples < 200 µm), the cartridges were opened using pliers and the filters were cut using sterilized blades in approximately 16 pieces (8 longitudinal and 1 horizontal cut). For this step, the cartridges were placed on sterilized aluminum foil on top of dry ice to prevent the material from defrosting. Filter pieces were then transferred using sterilized forceps (ethanol-flamed) into 1.5 ml Eppendorf tubes that contained the lysis buffer. Then, the protocol of the kit was followed; DNA was eluted to 50 µl and was stored at -80°C until amplification (typically within 1-5 days). Although DNA was extracted separately for the sediment trap particles < 200 µm and > 200 µm, these were pooled for subsequent analysis.

Once DNA was extracted, amplification was done by Polymerase Chain Reaction (PCR). The prokaryotic community was characterized by amplifying the V4-V5 regions of the 16S small subunit ribosomal RNA gene (SSU-rRNA) using primers 515F and 926R (Table S2). The eukaryotic community was characterized by amplifying the V9 region of the18S rRNA gene using primers 1389F and 1510R (Table S2). Primers contained the Illumina adaptors, the linker, and the barcoded indices. Amplification was done using the Q5 high-fidelity PCR kit (New England Biolabs) in a 25-µl reaction volume. The PCR thermal protocol consisted of an initial denaturation of 30 s at 98°C, 30 amplification cycles of 10 s at 98°C, 30 s at 56°C, and 30 s at 72°C, followed by a final extension of 2 min at 72°C, and a final holding of 4°C. The band size of the amplicons was visualized on a 1% agarose gel. Because sediment trap samples did not amplify during PCR, likely due to the organic matter present in sinking particles, we used the OneStep™ PCR Inhibitor Removal Kit (Zymo Research) following the manufacturer's instructions to remove the substances

4

that were inhibiting amplification. PCR was then carried out using 1 µl of diluted template (1:10). PCR products were purified using Agencourt AMPure Beads XP and the concentration was quantified using PicoGreen dsDNA Quantitation Reagent.

The PCR products were pooled in equimolar amounts (~ 10 ng µl^{-1}) in 1.5 ml Eppendorf tubes and were sequenced using a dual-barcode index on an Illumina MiSeq platform at the Institute for Genomic Medicine (IGM, University of California, San Diego). Demultiplexed raw reads were provided by IGM and have been deposited in NCBI under BioProject PRJNA445287 and Biosample accession numbers SAMN08784582– SAMN08784552 for 18SrRNA, and under BioProject PRJNA422420 and Biosample accession numbers SAMN08784494–SAMN08784464 for 16S rRNA.

Bioinformatic analyses

Initial quality control of the raw sequence reads (fastq files) was done using the workflow for read filtering, swarm Operational Taxonomic Unit (OTU) clustering, and taxonomic classification of the SSU-rRNA written by JP McCrow (https://github.com/allenlab/rRNA_pipeline). Briefly, paired-end reads were aligned using PEAR (Zhang *et al.,* 2014) and quality trimmed to Phred score 30 (Q30 minimum average in sliding window of size 2 bp) for 18S V9 amplicons or Q20 for 16S amplicons due to lower maximum quality scores. Possible chimeras were found and filtered using USEARCH (Edgar 2010). Amplicons were clustered using SWARM into OTUs (Mahé *et al.,* 2014) and taxonomic assignment was done by the best hit from GLSEARCH36 (Pearson, 2016) against the appropriate reference database: Silva v128 (Quast *et al.,* 2013) and phytoRef (Decelle *et al.,* 2015) were used for prokaryotes and potential plastid sequences, respectively, whereas PR2 with the taxonomic updates from the Tara Oceans-W2 was used for eukaryotes (de Vargas *et al.,* 2015).

The initial filtered OTU table for each library were processed further before multivariate analyses (Table S3). An additional qualitative control was done following the next steps: 1) For the 18S, non-eukaryote sequence reads were removed (archaeal, bacterial, organelle, and unassigned), whereas for the 16S, only bacterial and archaeal sequence reads were kept (eukaryotes were omitted). 2) OTUs with only 1 sequence read in the entire data set (singletons) were removed. 3) OTUs that were assigned to the same

5

genus, but that occurred multiple times in the OTU table were merged to have each genus only once (merge of over-split OTUs). 4) The taxonomy of each OTU down to the genus level was examined based on the information available in the World Register of Marine Species (WoRMS, www.marinespecies.org/) using the "Match taxa" function. This step helped to fill the gaps in the taxonomic information and to correct misassignations. 5) OTUs assigned only to a supergroup level (e.g., Opisthokonta_X, Stramenopiles_X) were removed. 6) OTUs assigned to the Class Streptophyta (land plants) were removed. These final OTU tables were used for analyzing the microbial communities of the sediment-trap and water-column samples. The treemap package in R (Tennekes, 2017) was used to described the most abundant microbes found in the water column and sediment traps.

Synechococcus *sequence analysis and classification*

The different *Synechococcus* strains were classified by sequencing the 16S-23S rRNA internal transcribed spacer (ITS) using primers ITS1F and ITS4R (Table S2). Amplification, sequencing, and bioinformatics processing (denoising, chimera detection, and OTU clustering) of *Synechococcus* sequences was carried out at RTL Genomics (Lubbock, Texas). The OTUs were assigned at a 97% cutoff. OTU classification was carried out using MOTHUR (Schloss *et al*., 2009) and a *Synechococcus* ITS database from Choi *et al*. (2014). OTUs that were assigned to the same strain were merged.

References

Bond, N.A., Cronin, M.F., Freeland, H., and Mantua, N. (2015) Causes and impacts of the 2014 warm anomaly in the NE Pacific. Geophys Res Lett **42:** 3414–3420.

Chavez, F.P., Pennington, J.T., Castro, C.G., Ryan, J.P., Michisaki, R.P., Schlining, B., *et al.* (2002) Biological and chemical consequences of the 1997–1998 El Niño in central California waters. Prog Oceanogr **54:** 205–232.

Choi, D.H., Hoon noh, J., and Lee, J.H. (2014) Application of pyrosequencing method for investigating the diversity of *Synechococcus* subcluster 5.1 in open ocean. Microbes Environ **29:** 17–22.

Decelle, J., Romac, S., Stern, R.F., Bendif, E.M., Zingone, A., Audic, S., *et al.* (2015) PhytoREF: a reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy. Mol Ecol Resour **15:** 1435–45

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares R., *et al*. (2015) Eukaryotic plankton diversity in the sunlit ocean. Science **348:** 1261605, DOI: 10.1126/science.1261605.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatic*s* **26:** 2460–2461

Fontanez, K.M., Eppley, J.M., Samo, T.J., Karl, D.M., and DeLong, E.F. (2015) Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. Front Microbiol **6:** 469. doi: 10.3389/fmicb.2015.00469.

Knauer, G.A., Martin, J.H., and Bruland, K.W. (1979) Fluxes of particulate carbon, nitrogen, and phosphorus in the upper water column of the northeast Pacific. Deep Sea Res I **26:** 97–108.

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2014) Swarm: robust and fast clustering method for amplicon-based studies. PeerJ **2:** e593.

McClathchie, S., Goericke, R., Leising, A., Auth, T., Bjorkstedt, E., Robertson, R., *et al*. (2016) State of the California Current 2015–16: Comparisons with the1997–98 El Niño. CalCOFI Rep **57:** 5–61.

Ohman, M.D., Barbeau, K., Franks, P.J.S., Goericke, R., Landry, M.R., and Miller, A.J. (2013) Ecological transitions in a coastal upwelling ecosystem. Oceanography **26:** 210–219.

Pearson, W.R. (2016) Finding protein and nucleotide similarities with FASTA. Curr Protoc Bioinform **53:** 3.9.1-3.9.25. doi: 10.1002/0471250953.bi0309s53.

Peterson, W.T., Keister, J.E., and Feinberg, L.R. (2002) The effects of the 1997–99 El Niño/La Niña events on hydrography and zooplankton off the central Oregon coast. Prog Oceanogr **54:** 381–398.

Peterson, W.T., Fisher, J.L., Strub, P.T., Du, X., Risien, C., Peterson, J., and Shaw, C.T. (2017) The pelagic ecosystem in the Northern California Current off Oregon during the 2014–2016 warm anomalies within the context of the past 20 years. J Geophys Res Oceans **122:** 7267–7290.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al*. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res **41:** D590–D596.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B. *et al*., (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol **75:** 7537–7541.

Stukel, M.R., Ohman, M.D., Benitez-Nelson, C.R., and Landry, M.R. (2013) Contributions of mesozooplankton to vertical carbon export in a coastal upwelling system. Mar Ecol Prog Ser **491:** 47–65.

Tennekes, M., and Ellis, P. (2017) treemap: Treemap Visualization. R package version 2.4-2.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014) PEAR: A fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics **30:** 614–620.

		Sediment traps			
Zone	**Dates drifting**	**Total (days)**	**Depth (m)**	**Treatment**	**Screened (um)**
Oligotrophic Offshore (OO) [Cycle 1]	22.April.2016 27.April.2016	5.6	100	Brine	200
			100	RNA later	200
			150	Brine	200
			150	RNA later	200
California Current (CC) [Cycle 2]	29.April. 2016 02.May.2016	3.3	97	Brine	200
			97	RNA later	200
			147	Brine	200
			147	RNA later	200
Transition Zone (TZ) [Cylce 3]	03.May.2016 06.May.2016	3.2	60	Brine	200
			60	RNA later	200
			150	Brine	200
			150	RNA later	200
Inshore (IN) [Cycle 4]	07.May.2016 10.May.2016	3.3	50	Brine	200
			50	RNA later	200
			150	Brine	200
			150	RNA later	200

		Water column			
Zone	**Date collected**		**Depth (m)**	**Volume (ml)**	**Screened (um)**
Oligotrophic Offshore (OO)	22.April.2016	-	50	650	500
	22.April.2016	-	100	650	500
	22.April.2016	-	150	650	500
California Current (CC)	02.May.2016	-	5	280	200
	30.April.2016	-	20	650	500
	30.April.2016	-	100	650	500
	30.April.2016	-	150	650	500
Transition Zone (TZ)	05.May.2016	-	12	280	200
	04.May.2016	-	20	280	200
	04.May.2016	-	60	650	500
	04.May.2016	-	150	650	500
Inshore (IN)	10.May.2016	-	7	650	500
	08.May.2016	-	12	280	200
	09.May.2016	-	12	280	200
	10.May.2016	-	50	650	500
	10.May.2016	-	150	650	500

Table S1. Summary information on samples collected to evaluate the microbial communities associated with sinking particles in the California Current Ecosystem. Dates drifting = dates of deployment and recovery of the sediment trap arrays in each zone. Total = time the trap array was drifting. Screened = pore size of the Nitex screen used to remove mesozooplankton before filtering the seawater and particles. The information in brackets refers to specific experimental cycles of CCE-LTER Process cruise P1604.

	Forward	Reverse	Amplicon size
Prokaryotes – 16S			
	515F	926R	
	GTGYCAGCMGCCGCGGTAA	CCGYCAATTYMTTTRAGTTT	~ 527 bp
Synechococcus			
	ITS1F	ITS4R	
	CTTGGTCATTTAGAGGAAGTAA	TCCTCCGCTTATTGATATGC	~ 350 bp
Eukaryotes – 18S V9			
	1389F	1510R	
	TTGTACACACCGCCC	CCTTCYGCAGGTTCACCTAC	~ 290 bp

Table S2. Primer information.

	16S	ITS	18S V9
Filtered reads	20,106	16,357	38,075
OTUs	1140	215	792

Table S3. Average number of sequence reads and number of Operational Taxonomic Units (OTUs) for all samples analyzed. The average was calculated from post-quality control data.
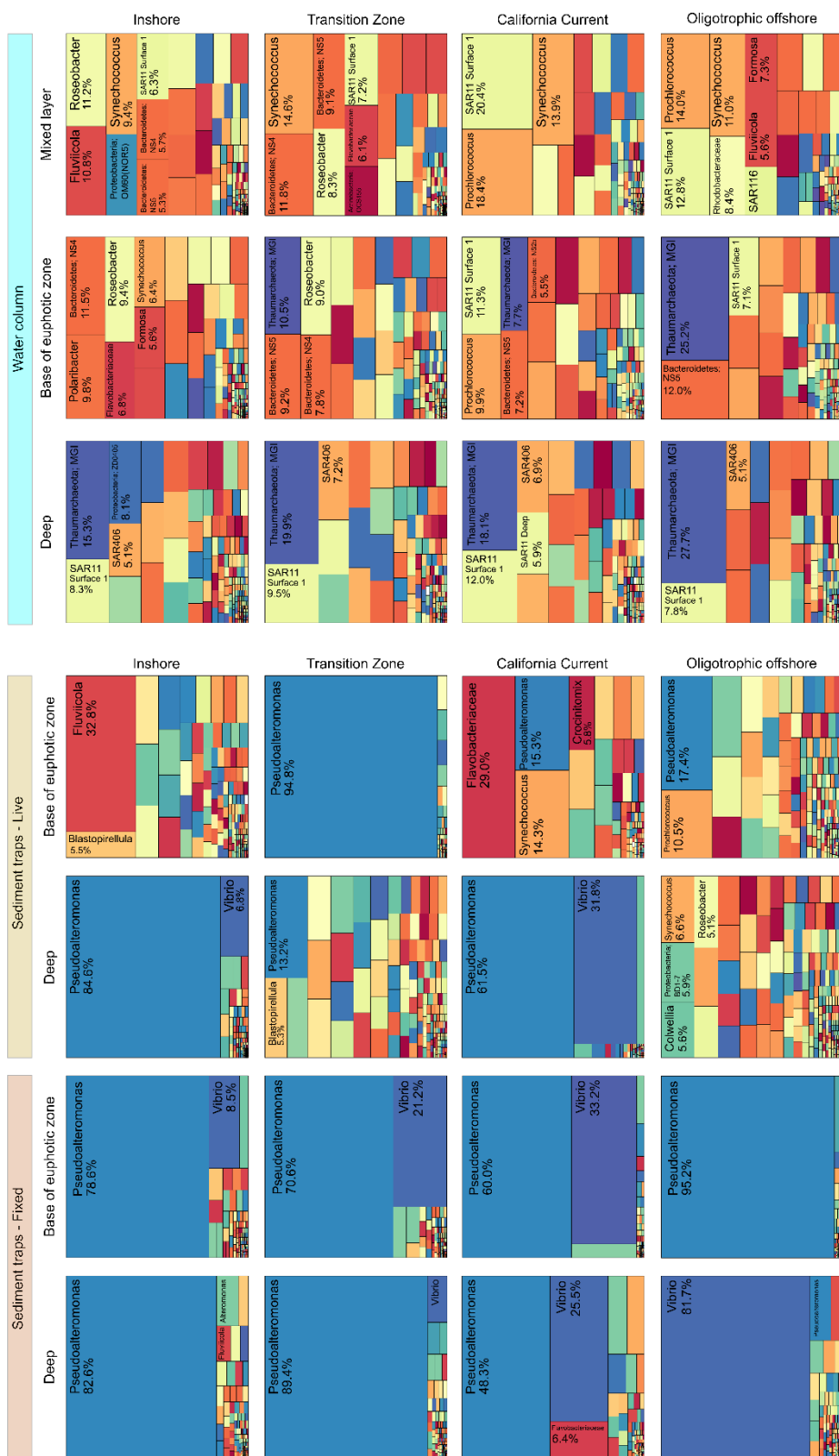
Figure S1. Percentages of protist sequences in the water column and on sinking particles in the California Current Ecosystem based on analysis of the 18S.

Figure S2. Percentages of bacteria sequences in the water column and on sinking particles in the California Current Ecosystem based on analysis of the 16S.
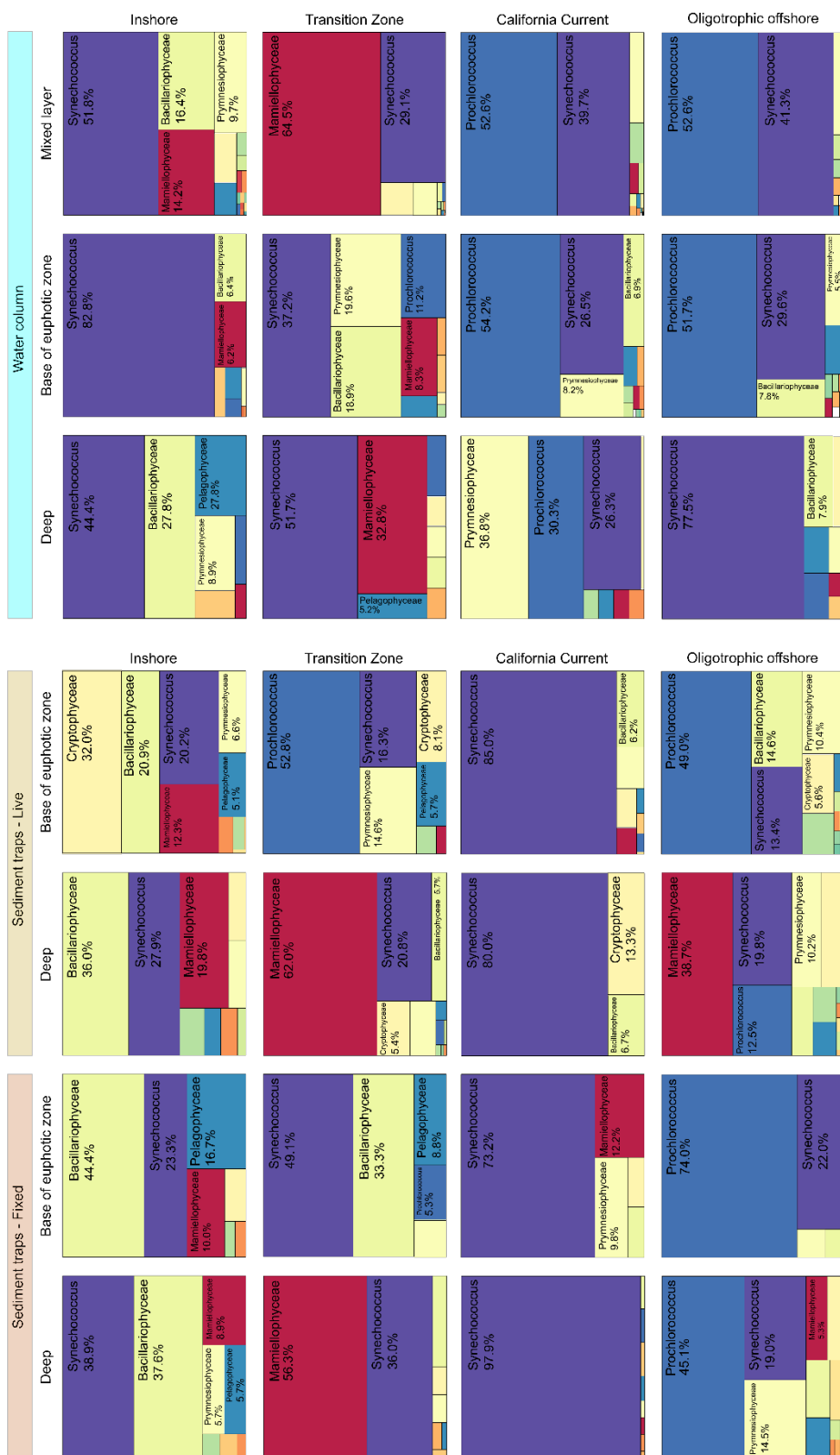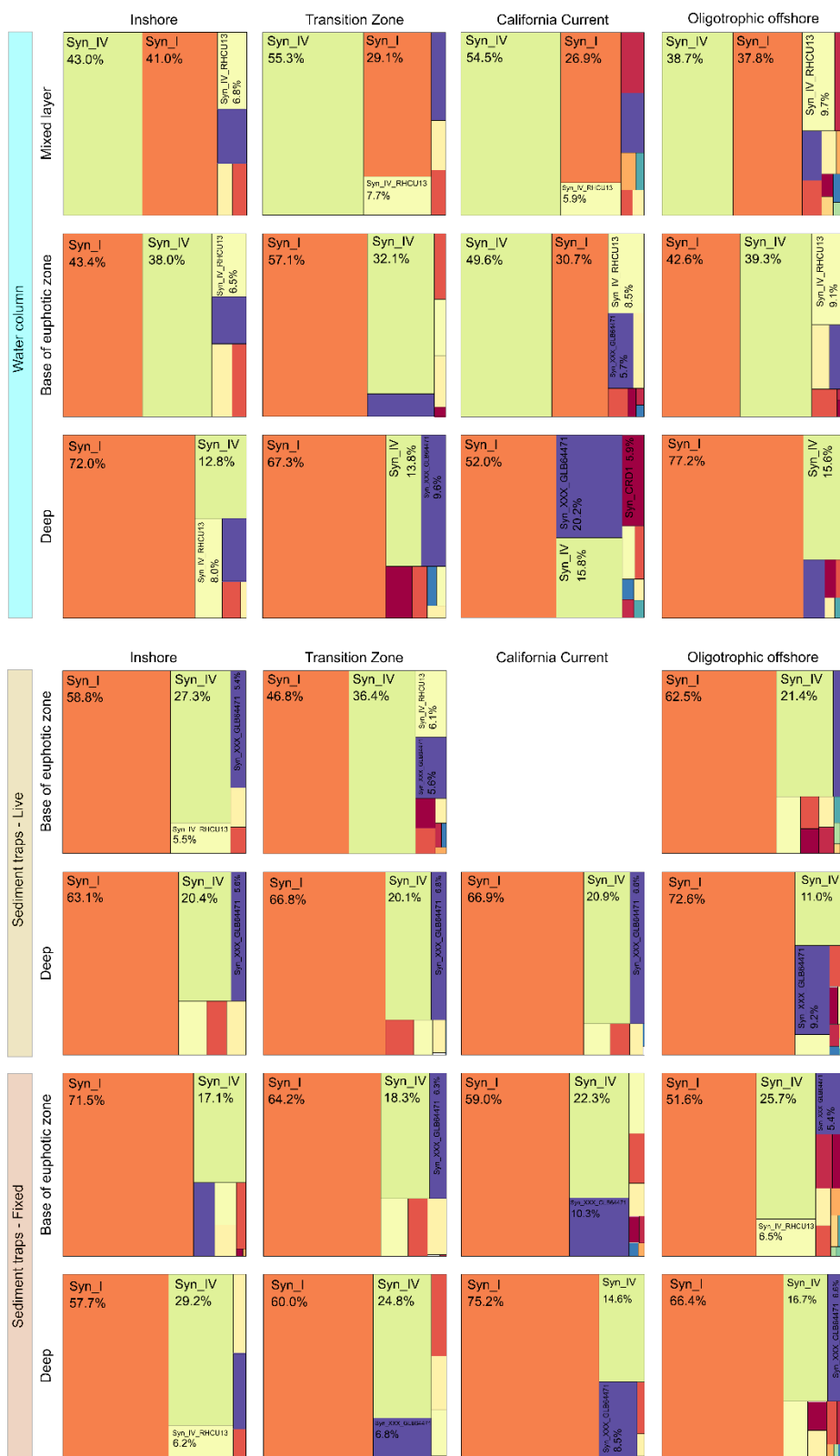
Figure S3. Percentages of plastid sequences in the water column and on sinking particles in the California Current Ecosystem based on analysis of the 16S.

Figure S4. Percentages of *Synechococcus* sequences in the water column and on sinking particles in the California Current Ecosystem based on analysis of the ITS.
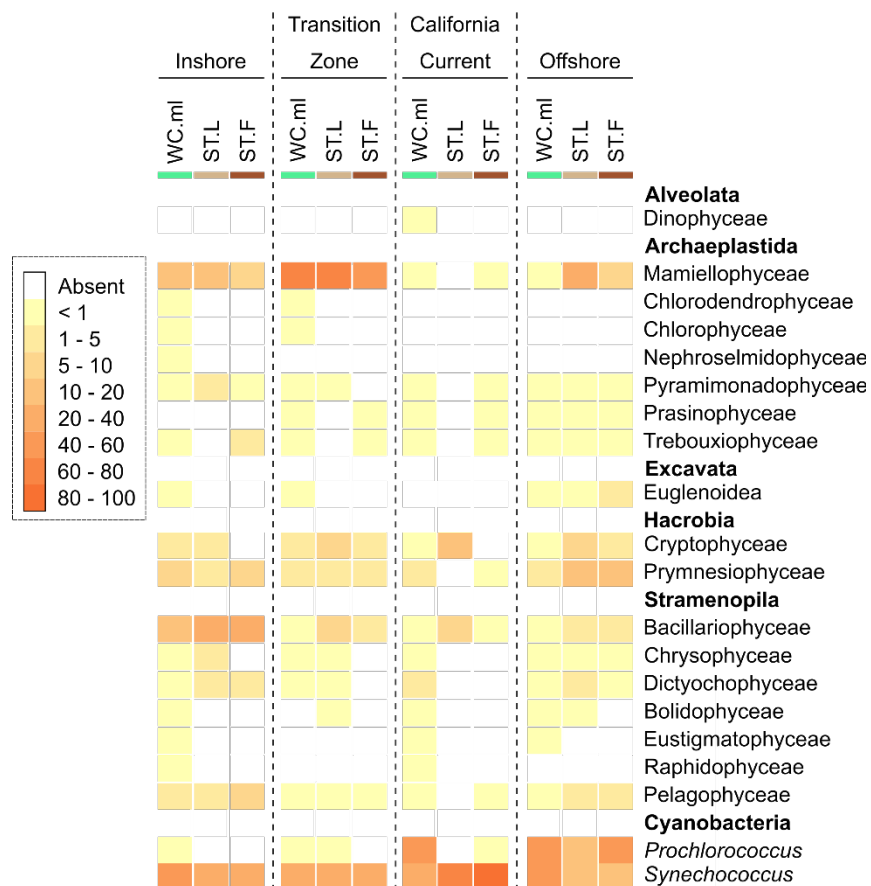
Figure S5. Percentages of eukaryotic phytoplankton and cyanobacteria (*Prochlorococcus* and *Synechococcus*) in the mixed layer and on sinking particles in the California Current Ecosystem obtained from plastid data amplified using the 16S rRNA. WC.ml: water column mixed layer, ST.L: live sediment traps, ST.F: fixed sediment traps.
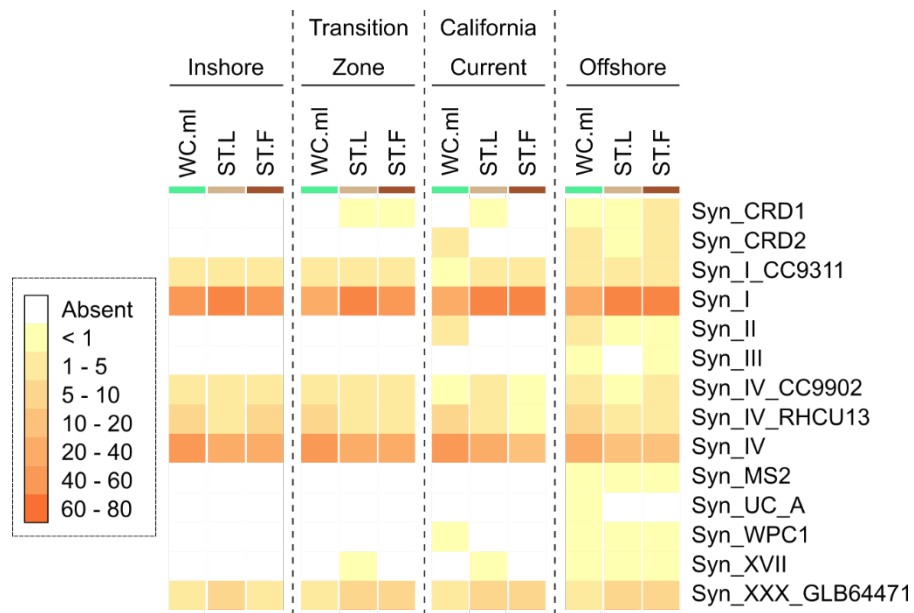
Figure S6. Percentages of *Synechococcus* strains in the mixed layer and on sinking particles in the California Current Ecosystem based on analysis of the ITS region. WC.ml: water column mixed layer, ST.L: live sediment traps, ST.F: fixed sediment traps.