Size fractionation informs microbial community composition and interactions in the eastern tropical North Pacific Ocean

Author names: Madeleine A. Thompson[1] (mthompson@seoe.sc.edu; https://orcid.org/0000-0003-1550-1505), David L. Valentine[2,3] (valentine@ucsb.edu; https://orcid.org/0000-0001-5914-9107), Xuefeng Peng[1,2] (xpeng@seoe.sc.edu; https://orcid.org/0000-0001-8696-1934)

Author affiliations: 1. School of Earth, Ocean, and Environment, University of South Carolina, Columbia, SC, USA, 2. Marine Science Institute, University of California, Santa Barbara, CA, USA, 3. Department of Earth Science, University of California, Santa Barbara, CA, USA

Author for correspondence: Xuefeng Peng, School of Earth, Ocean, and Environment, University of South Carolina, 701 Sumter Street, EWS 617, Columbia, SC, USA, 29201 Email: xpeng@seoe.sc.edu

Author Contribution Statement

Madeleine A. Thompson: Data Curation (equal), Formal Analysis (lead), Software (equal), Visualization (lead), Writing – Original Draft Preparation (lead), Writing – Review & Editing (equal). **David Valentine**: Funding Acquisition (equal), Resources (equal), Investigation (equal), Writing – Review & Editing (supporting). **Xuefeng Peng**: Conceptualization (lead), Data Curation (equal), Funding Acquisition (equal), Investigation (equal), Methodology (lead), Project Administration (lead), Resources (equal), Software (equal), Supervision (lead),

Validation (lead), Writing – Original Draft Preparation (supporting), Writing – Review & Editing (equal).

Abstract

Marine microorganisms are drivers of biogeochemical cycles in the world's oceans, including oxygen minimum zones (OMZs). Using a metabarcoding survey of the 16S rRNA gene, we investigated prokaryotic communities, as well as their potential interactions with fungi, at the coastal, offshore, and peripheral OMZ of the eastern tropical North Pacific. Water samples were collected along a vertical oxygen gradient, and large volumes were filtered through three size fractions, 0.22 μm, 2 μm, and 22 μm. The changes in community composition along the oxygen gradient were driven by Planctomycetota, Bacteroidota, Verrucomicrobiota, and Gammaproteobacteria; most are known degraders of marine polysaccharides and usually associated with the large particle-associated community. The relative abundance of Nitrososphaerota, Alphaproteobacteria, Actinomycetota, and Nitrospinota were high in free-living and small particle-associated communities. Network analyses identified putative interactions between fungi and prokaryotes in the particle-associated fractions, which have been largely overlooked in the ocean. In the small particle-associated network analysis, fungal ASVs had exclusively negative connections with SAR11 nodes. In the large particle-associated network analysis, fungal ASVs displayed both negative and positive connections with Pseudomonadota, SAR324, and Thermoplasmatota. Our findings demonstrate the utility of three-stage size-fractioned filtration in providing novel insights into marine microbial ecology.

Keywords: Oxygen Minimum Zone, metabarcoding, community composition, network analysis, fungi-bacteria interactions, size fractionation

Introduction

Microorganisms influence biodiversity and productivity in diverse marine environments. Oceanic oxygen minimum zones (OMZs), defined by oxygen concentration < 62.5 μM at intermediate depths (Stramma *et al.* 2008), form as a result of excess oxygen consumption by respiration and sluggish oxygen resupply via physical processes (Karstensen, Stramma and Visbeck 2008). The minimum oxygen concentration in certain OMZs, including the Arabian Sea and the eastern tropical North and South Pacific, is below detection (1 - 10 nM), and these OMZs are referred to as anoxic marine zones (AMZs) or anoxic OMZs (A-OMZs) (Ulloa *et al.* 2012; Canfield and Kraft 2022). Marine OMZs have expanded during the past 50 years, mainly due to anthropogenic influences, with continued expansion predicted for the coming century (Breitburg *et al.* 2018). Rising ocean temperatures, which decrease the solubility of oxygen in the water and can cause microbial blooms, are the leading cause of deoxygenation in marine ecosystems (Paerl and Otten 2013; Bertagnolli and Stewart 2018; Breitburg *et al.* 2018). Oceanic oxygen decline has reshaped the community composition of marine organisms, specifically prokaryotes (Bertagnolli and Stewart 2018).

The sharp vertical oxygen gradient found in OMZs provides an environment for ecological shifts between aerobic metabolisms (e.g., heterotrophic respiration, ammonium oxidation, and nitrite oxidation) and anaerobic metabolisms (e.g., denitrification and anaerobic ammonia oxidation (anammox)) (Canfield and Kraft 2022). Microorganisms have enzymatic and metabolic mechanisms for low-oxygen environments, for example, by deploying a higher ratio of high-affinity to low-affinity cytochrome oxidases (Kalvelage *et al.* 2015), allowing them to catalyze reactions that require oxygen under low oxygen levels (Peng *et al.* 2015, 2016, 2024) and

alternate substrates as terminal electron acceptors, such as nitrate (Bueno *et al.* 2012). Oxygen is the main driver of microbial community structure in oxygen minimum zones, influencing aerobic and anaerobic metabolic processes and biogeochemical cycling (Ulloa *et al.* 2012; Canfield and Kraft 2022). Additionally, temperature, salinity, and organic matter availability have been shown to determine the microbial community assembly in OMZs (Bryant *et al.* 2012; Beman and Carolan 2013; Peng, Jayakumar and Ward 2013; Rajpathak *et al.* 2018).

Organic matter is produced by different groups of phytoplankton in coastal oceans when compared to the open ocean, with larger eukaryotic primary producers contributing to a large proportion of primary productivity nearshore and Cyanobacteriota contributing to most of the primary production offshore (James *et al.* 2022). Coastal waters can contain more than double the chlorophyll a concentrations compared to offshore waters (Fuchsman *et al.* 2019). In certain regions of the open ocean, including OMZs, a deep chlorophyll a maximum (DCM) comprised of Cyanobacteriota with greater light-harvesting abilities (primarily *Prochlorococcus* and *Synechococcus*) is frequently observed (Lavin *et al.* 2010; Fuchsman *et al.* 2019; Callbeck *et al.* 2021). These differences in phytoplankton communities between nearshore and offshore seawater lead to differences in the vertical flux of organic matter (Fuchsman et al., 2019) and the formation of particulate organic matter, which influences the microbial community composition. Previous studies have found size fraction to be one of the strongest predictors of community composition (Ganesh *et al.* 2014; Suter *et al.* 2018). Particle-associated microorganisms attach to aggregates of organic matter and can create microhabitats enriched in nutrients and organic carbon (Ganesh *et al.* 2014). This allows a spectrum of aerobic and anaerobic metabolisms to occur in close proximity (Klawonn *et al.* 2015). Despite a wide range in the size of particles in

93 seawater (Davies *et al.* 2014), most studies in the ETSP and ETNP OMZs so far have used only
94 two size fractions, including one ~0.2 μm filter for the free-living size fraction and one filter with
95 a larger pore size for the particle-associated size fraction (Table 1). The selection of a small pore
96 size (e.g., 2 μm) for the particle-associated size fraction would result in the lumping of small and
97 large particles, while the use of only a 0.2 μm filter would result in the inclusion of all particles
98 on the 0.2 μm filter with the free-living size fraction, obfuscating the distinction between free-
99 living and particle-associated microbes.
100
101 In this study, we filtered large volumes (23 - 69 L, Supplementary Table S1) through three size
102 fractions (0.22 – 2.0 μm, 2.0 - 22 μm, and >22 μm) to separate both small (2.0 - 22 μm) and large
103 (>22 μm) particle-associated microbes from free-living (0.2 – 2.0 μm) microbial communities
104 from the eastern tropical North Pacific (ETNP) OMZ. By sequencing and analyzing the V4
105 region of the 16S rRNA gene from all three size fractions, we revealed previously overlooked
106 lineages that drive the differences in community composition between different size fractions
107 and under different oxygen levels. We also used network analyses that included previously
108 published fungal community composition data from the same samples to identify putative
109 keystone species that likely play an important role in OMZ biogeochemistry. We hypothesized
110 that a sharp redoxcline, consisting of the transitional zones between oxygenated and anoxic
111 waters, would have a selective effect on the microbial community composition and diversity. We
112 expected that all size fractions would exhibit strong community variations along redox gradients.
113 We hypothesized that these differences would be more pronounced at large particle-associated
114 size fractions compared with the small particle-associated and free-living size fractions because

of the variation in particle composition throughout the OMZ (Mestre *et al.* 2017; Maerz *et al.* 2020).

Materials and Methods

Sample Collection

Seawater samples (total of 16) were collected from the ETNP OMZ in surface/oxic (> 125 μM of O_2; 40 m at Station 1, 30 m at Station 2, and 10 m at Station 3), upper oxycline (above anoxic zone; 70 m and 83 m at Station1, 90 m at Station 2, and 33 m at Station 3), anoxic (<10 nM of O_2; 110 m and 120m at Station 1, 113 m and 120 m at Station 2, and 40 m, 45 m, 70 m at Station 3), and lower oxycline (below anoxic zone; 900 m at Station 2 and 1000 m at Station 3) (Figure 1) from March 22 to April 11, 2018, on *R/V Sally Ride* at three stations using 30-liter Niskin bottles (Figure 1) (Peng and Valentine 2021). When more than 30 L were filtered, water was collected in multiple Niskin bottles at the same depth on the same cast.

The dissolved oxygen concentrations and fluorescence in the water column were determined using a SBE 43 dissolved oxygen sensor (detection limit of 1μM) (Sea-Bird Scientific, Bellevue, WA, USA) and a Seapoint chlorophyll fluorometer (Seapoint et al., USA) attached to a conductivity, temperature, and depth (CTD) rosette. The mixed layer depth (MLD) was determined from vertical profiles of sigma-theta (σ_θ) calculated from temperature and salinity measured *in situ* during the CTD casts. The MLD criterion used was a threshold value of $\Delta\sigma_\theta$ = 0.03 kg m^{-3} from the value at 10 m (De Boyer Montégut 2004). Sulfanilamide and N(1-Naphthyl)ethylenediamine(NED) colorimetric spectrophotometer technique was used to measure ambient nitrite (Strickland and Parsons 1970). Nitrate profiles were measured using the

chemiluminescence method (Braman and Hendrix 1989). Lastly, the ortho-phthalaldehyde (OPA) method was used to measure ambient ammonium (Holmes et al., 1999; Taylor et al., 2007). For more details on nutrient concentration measurement methodology, refer to Tracey et al. (2022). To investigate microbial communities at different size fractions, the sampled seawater was sequentially filtered through a 47-mm diameter Whatman Grade 541 acid-hardened cellulose filter paper (22 μm nominal particle retention rating, GE Healthcare 1541–047, Marlborough, MA, USA), a 47-mm diameter polycarbonate filter (2.0 μm nominal pore size, Millipore Isopore TTTP-04700, Burlington, MA, USA), and a Sterivex filter (0.22 μm nominal pore size, Millipore SVGP01050, Burlington, MA, USA), using a peristaltic pump at a flow rate < 40 mL/min in a cold room (4°C) upon collection. For each sample set, 23 to 69 L of seawater was filtered (Supplementary Table S1). Microorganisms collected on the 22 μm filters were considered to represent the large particle-associated fraction (> 22 μm), microorganisms collected on the 2 μm filters were considered to represent the small particle-associated fraction (2 - 22 μm), and the microorganisms collected on the 0.22 μm filters represented the free-living fraction (0.22 - 2 μm). The filters were flash-frozen using liquid nitrogen before being stored at -80°C.

DNA Extraction

In the laboratory, DNA was extracted from the filters using the DNeasy Plant Mini Kit (Qiagen Cat. No. 69106), following the kit protocol except for the first step (cell lysis). The filter papers were cut into 2 mm squares using sterilized scissors. The filters were placed into 2 mL tubes containing 1 mL of 0.5 mm zirconia/silica beads (Biospec products #11079105z, Bartlesville, OK, USA), 600 mL of AP1 buffer, and 6 μL of RNase A. These tubes were placed in a Biospec

Mini-BeadBeater-16 for 90 seconds and incubated at 65 °C for 10 minutes. After incubation, the tubes were centrifuged at 20,000 g for 5 minutes, and the supernatant was transferred into new 2 mL tubes and neutralized with 195 µL of Buffer P3. The remaining extraction steps followed the DNeasy Plant Mini Kit protocols without further modification. The DNA quantity was determined using a Broad Range (BR) DNA Qubit fluorometer kit (ThermoFisher Scientific, Waltham, MA, USA), and the extracted samples were stored at -80 °C until amplicon library construction.

16S Amplicon Sequencing

The V4 region of the 16S rRNA gene was amplified using uniquely barcoded PCR primers 515F (5' - AATGATACGGCGACCACCGAGATCTACAC <i5> TATGGTAATT GT GTGCCAGCMGCCGCGGTAA - 3') and 806R (5' - CAAGCAGAAGACGGCATACGAGAT <i7> AGTCAGTCAG CC GGACTACHVGGGTWTCTAAT - 3'), where the <i5> and <i7> sequences are the 8-nt Illumina index sequences previously described (Kozich *et al.* 2013). Amplicon PCR reactions contained 1 µL of template DNA, 2 µL of 10 µM forward primer, 2 µL of 10 µM reverse primer, and 17 µL of AccuPrime Pfx SuperMix (ThermoFisher Scientific by Invitrogen). The thermal cycle settings started at 95°C for 2 min, followed by 30 cycles of 95 °C for 20 s, 55 °C for 15 s, and 72 °C for 5 min, and final elongation at 72 °C for 10 min. The concentrations of PCR products were normalized using the SequelPrep Normalization Kit (ThermoFisher Scientific by Invitrogen), cleaned using the DNA Clean and Concentrator kit (Zymo Research), visualized on an Agilent Tapestation, and quantified using a Qubit fluorometer. Samples were sequenced and demultiplexed at the University of California Davis Genome Center on the Illumina MiSeq platform with 250-nucleotide paired-end reads. A PCR-

184 grade water sample was included in extraction, amplification, and sequencing as a negative
185 control to assess DNA contamination.
186
187 ***Sequence Analysis***
188 Raw sequencing reads were analyzed with USEARCH (v11) to calculate the expected error
189 starting from 50 base pairs (bp) to 250 bp at 10 bp intervals (Edgar, 2010). This guided our
190 decision to trim the forward, and reverse reads to 150 bp. A total of 2,374,826 forward and
191 reverse reads were sequenced from 48 samples, 2,287,897 reads were merged, and 1,849,854
192 merged reads passed quality filtering, removing all sequences with a maximum expected error
193 greater than 1.0. The mean merged length was 254 bp (Supplementary Table S2). Removing the
194 low-quality tail of raw reads improved the quality and percentage of merged reads. Trimmed
195 forward and reverse reads were merged using USEARCH with the options "-fastq_maxdiffs 10".
196 Merged reads were quality filtered (with a maximum expected error of 1.0), dereplicated, and
197 denoised (using the command "usearch -unoise3") using USEARCH to obtain Amplicon
198 Sequence Variants (ASVs). 93.4% of merged reads were mapped to the ASVs to construct an
199 abundance table. The nucleotide sequences of ASVs were imported to QIIME2 v2022-11
200 (Caporaso *et al.* 2010) and classified by a Naïve Bayesian Classifier using the SILVA v138.1
201 99% identity OTUs from the 515F/806R region as the reference database (Quast *et al.* 2012).
202
203 ***Statistical and Network Analyses***
204 Environmental parameters, including salinity, temperature, the concentration of dissolved
205 oxygen, nitrate, nitrite, ammonium, and chlorophyll a at each station, were investigated using a
206 redundancy analysis (RDA) to determine their roles in shaping the microbial community

composition using the vegan package in R (Oksanen *et al.* 2022). Prior to RDA, a Chord transformation (decostand function in the vegan package) (Borcard, Gillet and Legendre 2011; Legendre and Borcard 2018) was performed for both the ASV abundance table and the environmental data table because they do not follow normality (determined by Shapiro-Wilk's tests in R) and homoskedasticity (determined by Breusch-Pagan tests in R). Variance inflation factors (VIF) were calculated, and co-varying environmental parameters with VIF values above five were removed from the RDA. A permutation test for RDA under a reduced model was run to determine the significance of the following environmental variables: temperature, the concentration of dissolved oxygen, nitrite, ammonium, and chlorophyll a. A permutational multivariate analysis of variance (PERMANOVA) and Hutchenson t-tests were run to determine the community variation across the oxygen gradient (surface/oxic, upper oxycline, anoxic, and lower oxycline).

Using the phyloseq package in R (McMurdie and Holmes 2013), the relative abundance of each ASV was calculated and combined with its taxonomic classification. The 1100 most abundant taxa across all samples were selected to generate stacked bar charts. The full ASV abundance table was rarefied to the same sequencing depth in each sample (12,135 sequences per sample, range of 569 – 76,742 reads per library) (Cameron *et al.* 2021). Using R, the rarefied ASV table was used to calculate the Shannon diversity (diversity function in vegan package), richness (specnumber function in vegan package), and evenness (Shannon diversity divided by the log of richness) of each sample. Using the ecolTest package in R, Hutchenson t-tests were completed to determine the significance of the diversity between the free-living, small particle-associated, and large particle-associated community at each station.

Differential abundance analysis of the ASVs was performed using the relative abundance of each ASV, which was normalized for composition bias using the Trimmed Mean of M-values (TMM) method in R (calcNormFactors function in edgeR package). The estimated dispersion was run on the normalized ASVs (estimateDisp function in edgeR package), and a differential abundance analysis was performed between oxic, upper oxycline/anoxic, and lower oxycline regions (makeContrasts function in limma package). A Quasi-Likelihood F Test was performed on the output of the differential abundance analysis (glmQLFTest function in edgeR package), and then the counts per million (output from the expression analysis) for each 12ASVs that are differentially regulated were counted (decideTestsDGE function in edgeR package). The counts per million for the top 48 ASVs were merged with the taxonomy. The 48 most abundant ASVs that were differentially abundant across conditions were visualized using ggplot2.

Network analyses of the prokaryotic and fungal (Peng and Valentine 2021) microbial communities at each size fraction were performed using the Molecular Ecological Network Analyses Pipeline (MENAP) (Zhou *et al.* 2010, 2011; Deng *et al.* 2012; Xiao *et al.* 2022). Peng and Valentine (2021) used an ITS2 primer set on the same samples in this study to identify fungal ASVs in the ETNP OMZ. The rarefied ASV table was used to construct the network using the random matrix theory-based network approach (Chavda, Deota and Kota 2014). For all size fractions, only ASVs observed in more than 68.75% of all samples (11 samples out of 16 samples) were included in each network analysis. This percentage was chosen to include enough ASVs to form the network while being stringent to target the predominant ASVs. The relative abundance of ASVs sequences was transformed using a centered log ratio and a Pearson

253 correlation coefficient (Zhou *et al.* 2010). The correlation and significance were determined
254 using a Chi-square test on Poisson distribution. Network global properties, individual nodes'
255 centrality, and module separation and modularity calculations (greedy modularity optimization)
256 were calculated as part of the MENAP pipeline. The completed network was visualized in
257 Cytoscape v3.9.1 (Shannon *et al.* 2003). Putative keystone species were identified using the
258 criteria $P_i < 0.6$ and $Z_i > 1.5$ for module hubs and $P_i > 0.6$ and $Z_i > 1.5$ for network hubs. Z_i
259 represents how well a node is connected to other nodes in the same module. P_i represents how
260 well a node is connected to different modules, calculated using equations from Guimerà &
261 Amaral (2005) (Xiao *et al.* 2022). The keystone identification was putative because these ASVs
262 have not been experimentally verified as keystone species. The MENAP used included a recently
263 added module, iDIRECT, which removes indirect connections between nodes by eliminating
264 self-looping and the values of the total interaction strengths outside their natural range (Xiao *et*
265 *al.* 2022). Chloroplast DNA identified in the network analysis was further classified using NCBI
266 blastn.
267
268 **Results**
269 ***Environmental context***
270 At station 1 (the peripheral OMZ), oxygen concentrations were approximately 200 μM in the
271 mixed layer (down to 50 meters, Figure 2A). Then, oxygen sharply decreased in the oxycline
272 (between 50 and 80 meters) and remained below detection starting at ~120m (Figure 2A). In the
273 surface, chlorophyll a concentrations ranged between 0.4 and 0.8 mg m^{-3}, and displayed a
274 subsurface maximum (4 mg m^{-3}) at the bottom of the mixed layer (Figure 2A). Below a local
275 chlorophyll a concentration minimum at about 2 mg m^{-3} between 60 and 70 meters, there was a

deep chlorophyll maximum (DCM, 2.5 mg m^{-3}) near the oxic-anoxic interface (~ 85 m). Ammonium concentrations displayed a subsurface maximum of 90 nM in the oxycline and another peak at 90 m (55 nM; Figure 2D). A large primary nitrite maximum reaching 1.5 μM was observed in the oxycline, and the nitrite buildup in the anoxic depths was ~0.1 μM. Lastly, nitrate was depleted in the surface and increased in the oxycline and anoxic zone and remained at approximately 29 μM for the remaining sampling depths (Figure 2D).

At Station 2 (the offshore OMZ), the overall trend of the oxygen and chlorophyll a concentration depth profiles resembled those at Station 1, while both the shallow and deep chlorophyll maxima were positioned slightly deeper at Station 2 than at Station 1 (Figure 2B). A defining feature of permanent OMZs, the secondary nitrite maximum (SNM) (Codispoti *et al.* 1986; Peng *et al.* 2015; Kelly *et al.* 2021; Travis *et al.* 2023), was observed at around 250 m at the offshore OMZ, reaching 1.6 μM of nitrite (Figure 2E). In contrast, the SNM was absent in the peripheral OMZ. Ammonium concentration at Station 2 peaked twice above the oxycline (first peak of 52.0 nM and second peak of 50.0 nM). At this station, one sample was collected from the lower oxycline at 900 m, where oxygen concentration was approximately 3 μM (Figure 2B).

At Station 3 (the coastal OMZ), there was a sharp oxycline between the surface and 40 m (Figure 2C). Oxygen remained below detection (10 nM) from ~40 m to 840 m. Oxygen increased to 12 μM at 1000 m in the lower oxycline. In the mixed layer (from surface to 11 m), chlorophyll a concentrations were about three times as high as those at Stations 1 and 2, reaching a shallow maximum of 11.7 mg m^{-3}. The chlorophyll a concentration remained above 2 mg m^{-3} down to 70 m, which intersected with the upper part of the anoxic depths (Figure 2C). The surface maximum

299 of ammonium concentration reached 525.6 nM at 10 m; below 25 m the ammonium
300 concentrations remained low (< 15 nM). The primary nitrite maximum (0.7 μM) was positioned
301 below the subsurface ammonium peak at 20 m, and the SNM (2.8 μM) was observed at 160 m
302 (Figure 2F).
303
304 ***Prokaryotic community composition and diversity***
305 Sequences of the V4 region of the 16S rRNA gene from the ETNP OMZ comprised 5,050 ASVs,
306 with 4,646 ASVs affiliated with Bacteria and 404 ASVs affiliated with Archaea (Figure 3).
307 93.4% of reads were mapped to all ASVs (Supplementary Table S2). Across all samples,
308 bacteria accounted for 78.5% of the total microbial community. Archaea accounted for 21.5% of
309 the total microbial community (Supplementary Table S3). The dominant bacterial phyla included
310 Pseudomonadota (including the classes Alphaproteobacteria and Gammaproteobacteria),
311 Cyanobacteriota, and Bacteroidota, comprising 28.5% (18.1% are Alphaproteobacteria and
312 10.4% are Gammaproteobacteria), 13.4%, and 12.1% of the prokaryotic community,
313 respectively. The dominant archaeal phyla consisted of Nitrososphaerota, Thermoplasmatota,
314 and Nanoarchaeota, comprising 11.2%, 9.7%, and 0.6% of the prokaryotic community,
315 respectively (Supplementary Table S3). The 1100 most abundant ASVs included in Figure 3
316 accounted for approximately 90.5% of all merged reads (2,287,897 in total).
317
318 The prokaryotic community composition and diversity were distinct between different size
319 fractions (Supplementary Table S4). Large particle-associated communities were enriched with
320 Planctomycetota (up to 27%), Bacteroidota (up to 53%), Verrucomicrobiota (up to 16%), and
321 Gammaproteobacteria (up to 24%) (Figure 3, Supplementary Figures S1 – S4, Supplementary

Table S4). In contrast, the relative abundance of Nitrososphaerota (up to 51%), Alphaproteobacteria (up to 43%) (Supplementary Figure S5), Actinomycetota (up to 14%), and Nitrospinota (up to 14%) were high in free-living and small particle-associated size fractions (Figure 3, Supplementary Table S4). The relative abundance of the NB1-j group (formerly a member of Deltaproteobacteria) was high (up to 3.8%, Supplementary Table S4) in the small particle-associated and large particle-associated communities at the oxycline and anoxic depths (83 to 1000 m).

Alpha diversity

The Shannon diversity and evenness were similar between the free-living and small particle-associated communities at the peripheral OMZ (mean of 0.43 and 0.35, respectively). In contrast, the prokaryotic diversity of the large particle-associated community was low (mean of 0.23) (Supplementary Figure S6). In the offshore OMZ, the Shannon diversity of the small particle-associated and large particle-associated size fractions (mean of 0.14 and 0.20, respectively) was low compared to the free-living community (mean of 0.46) (Supplementary Figure S6). There was a significant difference (Hutcheson's t-test, $p < 0.05$) in diversity between large particle-associated and free-living microbial communities at Stations 1 and 2 (Supplementary Table S5). There was a significant difference (Hutcheson's t-test, $p < 0.05$) in diversity between the small particle-associated and large particle-associated microbial communities at all stations sampled (Supplementary Table S5). There was also a significant difference ($p < 0.05$) between the free-living and small particle-associated at the offshore and coastal OMZ stations (Supplementary Table S5). Overall, the diversity of the free-living communities (mean of 1.30) was the highest

344 compared to that of small particle-associated and large particle-associated communities (mean of
345 0.90 and 1.01, respectively).
346
347 ***Microbial community composition along the vertical oxycline***
348 The community composition of bacteria and archaea was significantly different between the
349 different oxygen regimes (surface/oxic (O_2 >150 μM), upper oxycline (detection limit < O_2 <
350 150 μM), anoxic (O_2 < detection limit), and lower oxycline (O_2 < detection limit and below
351 anoxic depths), PERMANOVA p-value = 0.0001, Supplementary Table S6). Temperature, the
352 concentration of dissolved O_2, nitrite (NO_2^-), and chlorophyll a significantly correlated with the
353 prokaryotic community composition (Figure 4, Supplementary Table S7, Supplementary Figure
354 S7, Supplementary Figure S8).
355
356 ASVs from Flavobacteriales, Planctomycetota, and Verrucomicrobiota were the key members
357 that drove the differences in microbial community composition between different oxygen
358 regimes (Figure 5, Supplementary Figure S9). The NS4 and NS5 marine groups
359 (Flavobacteriales), *Formosa* (Flavobacteriales), *Pirellula* (Planctomycetota), *Lentimonas*
360 (Verrucomicrobiota), Kiritimatiellaceae (Verrucomicrobiota), and Marine Group II Archaea
361 were enriched (up to 3.7%, 1.6%, 2.2%, 2.7%, 1.7%, 0.2%, and 10.9%, respectively) in the
362 surface/oxic communities. In contrast, *Crocinitomix* (Flavobacteriales), Rhodobacteraceae
363 (Alphaproteobacteria), NB1-j, Desulfobacterota, and Woesearchaeales (Nanoarchaeota) were
364 enriched (up to 1.1%, 1.9%, 2.0%, 2.3%, and 2.2%, respectively) in the upper oxycline and
365 anoxic depth communities. Three lineages, including NS9 marine group (Flavobacteriales),
366 Phycisphaeraceae (Planctomycetota) and *Roseibacillus* (Verrucomicrobiota) included ASVs

367 enriched in the surface/oxic communities and ASVs enriched at low-oxygen (< 10 μM) depths.
368 One Methylomondaceae (Gammaproteobacteria) ASV and one OM190 Group (Planctomycetota)
369 ASV were only detected in the lower oxycline (900 or 1000 m depth). In the coastal OMZ,
370 where the depths of the oxycline (10 - 35 m) and the anoxic layer were shallow (starting at ~35
371 m, Figure 2), the prokaryotic community at the oxycline and anoxic depths included many ASVs
372 also abundant at surface depths (Figure 5). Most of the lineages (34 out of 46) driving the
373 difference in community composition between depths were higher in abundance in large and/or
374 small particle-associated communities (Supplementary Figure S9).
375
376 ***Co-occurrence network analyses***
377 MENAP constructed co-occurrence networks for each of the three size classes: free-living, small
378 particle-associated, and large particle-associated, consisting of various nodes (ASVs), edges
379 (links), and the percentage of positive edges (Table 2).
380
381 There are 13 modules present in the free-living network, which consists of ASVs from 12
382 different taxa (Figure 6, Supplementary Table S8). Alphaproteobacteria accounted for 41% of all
383 ASVs in the free-living network, with SAR11 accounting for approximately 1/3 of all ASVs in
384 the network. Eight ASVs were identified as putative keystone species ($Z_i > 1.5$ and $P_i < 0.6$),
385 including SAR11 clade I (Alphaproteobacteria, ASV6), SAR11 clade II (Alphaproteobacteria,
386 ASV315 and ASV488), SAR86 clade (Gammaproteobacteria, ASV43 and ASV194),
387 *Prochlorococcus* (Cyanobacteriota, ASV3), *Nitrospina* (Nitrospinota, ASV221), and MB11C04
388 marine group (Verrucomicrobiota, ASV857).
389

Seven modules were identified in the small particle-associated network, which included ten prokaryotic taxa and two fungal phyla (Ascomycota and Basidiomycota) (Figure 7, Supplementary Table S8). Alphaproteobacteria accounted for 1/3 of all ASVs in the small particle-associated network, with SAR11 accounting for approximately 31% of all ASVs present in the network, with the remaining Alphaproteobacteria belonging to Rhodospirillales. *Rhodotorula* (Basidiomycota, ASV-F1) and *Cladosporium* (Ascomycota, ASV-F74) had exclusively negative connections with prokaryotic nodes but were positively connected to each other (Figure 7). Five ASVs were identified as putative keystone species ($Z_i > 1.5$ and $P_i < 0.6$), including SAR11 clade II (Alphaproteobacteria, ASV18 and ASV460), *Synechococcus* (Cyanobacteriota, ASV13), Chloroplast DNA (phototrophic eukaryotes, ASV74), and *Cladosporium* (Ascomycota, ASV-F74).

Two modules were identified in the large particle-associated network, which included ten prokaryotic taxa and one fungal phylum (Basidiomycota) (Figure 8, Supplementary Table S8). Gammaproteobacteria accounted for 24% of all ASVs in the large particle-associated network, with both Vibrionales and Alteromonadales each accounting for approximately 9% of all ASVs present in the network, with the remaining 6% Gammaproteobacteria belonging to Enterobacterales, Oceanospirillales, and Cellvibrionales. Five ASVs were identified as putative keystone species ($Z_i > 1.5$ and $P_i < 0.6$), including Vibrionaceae (Gammaproteobacteria, ASV130), SAR11 clade I (Alphaproteobacteria, ASV10), *Winogradskyella* (Bacteroidota, ASV675), and Phycisphaeraceae (Planctomycetota, ASV468 and ASV388).

Discussion

Experimental biases

The large volume seawater filtration in this study (23 - 69 L, Supplementary Table S1) is typically not adopted by conventional methods without any size fractionation because the 0.22 μm filters would be clogged. The large volume of seawater filtered raised the probability for free-living microorganisms to be trapped with large particle-associated communities due to effective filter pore size reduction. However, the low relative abundance of known free-living taxa such as Nitrososphaerota in the large particle-associated communities (Figure 3, Supplementary Figure S10) indicates that there was minimal "cross-talking" or interference/mixing between filter sizes. On the other hand, the low relative abundance of known particle-associated taxa, such as Planctomycetota and Verrucomicrobiota in the free-living communities (Figure 3), suggests that the gentle filtration method we used minimized particle disintegration.

While the slow flow rate we used during sequential filtration was intended to keep the integrity of the particles, it resulted in long filtration duration for large volumes. At a flow rate of ~ 40 mL/min, it took between 10 and 29 hours to filter 23 – 69 L of seawater (Supplementary Table S1). Although the microbes were caught on one of the filters undergoing filtration at 4°C, which is an environment very different from *in situ* conditions, they could still grow in population size. While microbial growth rates vary depending on the lineage, most of them displayed a maximum growth rate less than 1 d^{-1} at temperatures more than 10 °C higher than the filtration temperature we used in the field (Long *et al.* 2021). If the growth rate of microbes captured on filters scale with their maximum observed growth rate, then the community composition from our sequencing analysis may overrepresent fast-growing lineages such as Gammaproteobacteria (e.g.

436 *Vibrionaceae*). In the four samples collected from anoxic depths, obligate anaerobes, such as
437 Candidatus *Methylomirabilis oxyfera* NC10, were not expected to grow as the filtration was
438 conducted in air, and therefore they might be slightly underrepresented compared to facultative
439 organisms.
440
441 Additionally, PCR protocols are known to introduce bias in metabarcoding studies (Silverman *et*
442 *al.* 2021). Primers introduce some bias during the amplification step of PCR due to mismatches
443 between the primer and the corresponding binding site (Read and Pietersen 2016). The primers
444 used in this study (515F/806R) were from the Earth Microbiome Project (EMP) (Caporaso *et al.*
445 2012, 2018). However, a recent study found that the ubiquitous and abundant SAR11 clade and
446 Marine Group I Nitrososphaerota were typically underestimated by the EMP primer pair (Parada
447 et al. 2016). Therefore, we are aware that our results may underestimate the relative abundance
448 of the SAR11 clade and Marine Group I Nitrososphaerota. Despite this caveat, both the SAR11
449 clade and Marine Group I Nitrososphaerota were one of the most abundant taxa in all samples,
450 accounting for up to 30% of the total prokaryotic community.
451
452 ***Large volume filtration reveals previously overlooked large particle-associated communities***
453 Most of the 46 most abundant ASVs that displayed significant variation in relative abundance
454 along the oxycline were enriched in the large particle-associated communities (Figure 5). Many
455 of these 46 ASVs, such as Verrucomicrobiota, Bacteroidota, and Planctomycetota, although
456 previously reported in OMZs (Ganesh et al., 2015; Suter et al., 2018; Torres-Beltrán et al., 2019),
457 were not recognized as important drivers of community composition. We attribute these novel

458 insights in part to the application of the three-stage filtration we used to separate large particle-
459 associated from small particle-associated and free-living communities.
460
461 In a previous study conducted at a coastal ETNP OMZ station near our coastal OMZ station that
462 also used three size fractions (0.2, 1.6, and 30 μm) and the same EMP primer set (Ganesh et al.,
463 2015), the relative abundance of Verrucomicrobiota, Bacteroidota, and Planctomycetota were
464 lower than that in our coastal OMZ station while the relative abundance of Marinimicrobia
465 (SAR406) were higher than that in our study. This difference was likely due to the difference in
466 particle concentration in the water column, as the chlorophyll a concentration sampled by
467 Ganesh and colleagues in June was an order of magnitude lower than that at the coastal OMZ in
468 our study, which took place in April. The greater chlorophyll a concentration in our study likely
469 resulted in much higher abundance of particles that supported the particle-associated lineages.
470 Similarly, Station 2 in another study in the ETNP by Beman and colleagues (2021) was very
471 close to our offshore OMZ Station, which were both sampled in April, but the maximum
472 chlorophyll a concentration at our station was about three times as high as that in Beman et al.
473 (2021). Typical particle-associated taxa such as Verrucomicrobiota and Planctomycetota were
474 not even detected by Beman and colleagues (2021). This was unexpected because the maximum
475 chlorophyll a concentration at Station 2 of the Beman et al. (2021) study was twice as high as
476 what was reported for the coastal OMZ by Ganesh et al. (2015), so it is likely that the
477 concentration of particulate organic matter was higher in the Beman et al. (2021) study than in
478 the Ganesh et al. (2015) study. We interpret the lack of Verrucomicrobiota and Planctomycetota
479 representation in the Beman et al. (2021) study as a result of "dilution effect" when particle-
480 associated microbial communities were sampled along with free-living communities. The

application of three size fractions during filtration is necessary to reveal lineages associated with particles, which are hotspots for chemical transformations (Wright et al. 2012), and large phytoplankton such as diatoms that are known to host bacterial epibionts (Crenn, Duffieux and Jeanthon 2018).

The prokaryotic community composition in the water column redoxcline of the Cariaco Basin (8 or 12 L of seawater) partitioned into > 2.7 μm and 0.2 - 2.7 μm fractions found Verrucomicrobiota, Planctomycetota, Bacteroidota, and Alteromonadales to be enriched in the > 2.7 μm communities (Suter *et al.* 2018). Our study found the same lineages to be particle-associated, but particularly in the large particle-associated (> 22 μm) communities and less abundant in the small particle-associated (2 - 22 μm) communities (Figure 5, Supplementary Figures S1 – S4). Moreover, we found that the relative abundance of Verrucomicrobiota can reach 5 – 15% in large particle-associated and small particle-associated communities, whereas in the Cariaco Basin, Verrucomicrobiota was only 1 – 5% of the particle-associated (> 2.7 μm) community (Suter *et al.* 2018). Previous studies have found that Verrucomicrobiota, specifically MB11C04 marine group and Puniceicoccaceae, both of which are also found in our study, are specialist consumers of sulfated methyl pentose during diatom blooms (Orellana *et al.* 2022). This suggests that Verrucomicrobiota were particularly enriched on large (> 22 μm) particles, and filtration methods without including a large filter may underestimate their importance on large particles.

Prokaryotic community composition shifts along the redoxcline

The redoxcline in oceanic OMZs is known to shape microbial community assembly (Ulloa *et al.* 2012; Bertagnolli and Stewart 2018). This has been reported for the ETNP OMZ (Faull *et al.* 2020; Beman *et al.* 2021), and a recurring microbial community shift was observed across the transitional layers among all major oceanic OMZs, including the ETNP, the eastern tropical South Pacific (ETSP) (Ganesh *et al.* 2014), and the Arabian Sea (Fernandes, Shenoy and Damare 2020). Not only are our findings consistent with these previous reports, but we also identified the most abundant members that contributed to the distinction between the prokaryotic communities along the redoxcline (Figure 5). In this section, we highlight three taxa (Flavobacteriales, Verrucomicrobiota, and Planctomycetota) that demonstrated the strongest shift in their relative abundance along the redoxcline.

Flavobacteriales

Different lineages of Flavobacteriales were dominant along the ETNP redoxcline. *Formosa*, NS4, NS5, NS2b, and NS9 marine groups were primarily present and abundant in the mixed layer and largely absent in the oxycline and anoxic depths (Supplementary Figure S2) consistent with previous studies of the ETNP OMZ (Faull *et al.* 2020; Beman *et al.* 2021; Pajares 2021). Although uncultivated, *Formosa* and the NS4, NS5, NS2b, and NS9 marine groups of the Flavobacteriales are known degraders of large phytoplankton-derived biomass in the ocean (Teeling *et al.* 2016; Unfried *et al.* 2018). In all stations sampled, Flavobacteriales (except for NS5 marine group) were enriched in the large particle-associated size fraction, suggesting that the increased prevalence of these Flavobacteriales might correspond with the accessibility to a

source of larger phytoplankton biomass or zooplankton fecal pellets they typically degrade (Turner 2002), which were typically associated with the primary chlorophyll maximum.

Conversely, *Crocinitomix* was found nearly exclusively at oxycline/anoxic depths (Supplementary Figure S2). *Crocinitomix* has not been reported in previous studies conducted at the ETNP, ETSP, and the Arabian Sea OMZs. *Crocinitomix* cultivars are strict aerobes with restricted catabolic repertoire (Bowman 2015), so it was surprising to observe their apparent niche in the upper oxycline depths of the ETNP water column, as well as anoxic depths at the coastal station (Supplementary Figure S2). A recent study in the Namibian OMZ using DNA stable isotope probing found that *Crocinitomix* assimilated ^{13}C-labeled diatom exopolysaccharides at low-oxygen depths (down to ~20 μM O_2) and even in the anoxic sediments (Vuillemin, Coskun and Orsi 2022). These findings indicate a Flavobacteriales lineage previously overlooked in algal polysaccharide degradation in the absence of oxygen.

Verrucomicrobiota

Like Flavobacteriales, most lineages of Verrucomicrobiota were present and abundant in the surface ocean. However, three ASVs from the genus *Roseibacillus* were largely absent in the mixed layer and displayed a subsurface maximum near the oxic-anoxic interface (Supplementary Figure S3), which was near the maximum concentrations of N_2O (Supplementary Figure S8; Kelly et al. 2021). *Roseibacillus* was the most abundant and prevalent among the four Verrucomicrobiota genera found in the ETNP (Supplementary Figure S3). The *Roseibacillus* ASV7 was found in the top ten most abundant ASVs in all large particle-associated communities in the ETNP OMZ (Supplementary Table S9). In the large particle-associated communities,

network analysis identified many positive correlations between ASV7 and Cyanobacteriota (e.g. *Prochlorococcus* ASV5 and *Synechococcus* ASV13 and ASV59) and negative correlations between ASV7 and SAR11 (e.g. ASV12 and ASV40). This suggests *Roseibacillus* may benefit from carbon fixed by Cyanobacteriota while they may be competing for substrates with free-living heterotrophic bacteria such as SAR11. This is consistent with a prior hypothesis from a study using stable isotope probing that *Roseibacillus* acquire C and N by transporting the oligopeptides released through protein hydrolysis from extracellular enzymes manufactured by other bacteria (Orsi *et al.* 2016).

The coincidence of the subsurface maximum in the relative abundance of *Roseibacillus* and the maximum concentrations of N_2O (Kelly et al. 2021, Supplementary Figure S8) suggests a potential contribution by *Roseibacillus* to the formation of the subsurface N_2O maximum. The metagenome-assembled genomes of *Roseibacillus* from the ocean all possess the nitric oxide reductase gene (*norB*) (Tully, Graham and Heidelberg 2018; Chen *et al.* 2023; Mukherjee *et al.* 2023). We postulate that *Roseibacillus* contributes to N_2O production in the ETNP OMZ, fueled by carbon produced by Cyanobacteriota in the DCM and nitric oxide produced by nitrite-reducing bacteria such as Rhodospirillales and *Marinobacter* (Bandekar et al., 2018). Future studies could test this hypothesis by measuring the abundance of *Roseibacillus norB* transcript and protein abundance using metatranscriptomics and metaproteomics.

Planctomycetota

In the ETNP OMZ, Planctomycetota were primarily associated with large particles, consistent with previous reports from low-oxygen regions of the ocean (Suter *et al.* 2018; Torres-Beltrán *et*

570 *al.* 2019). The two most abundant Planctomycetota taxa, OM190 group and Phycisphaerales
571 (Supplementary Figure S1), are known to be associated with macroalgae and capable of
572 degrading algal sulfated polysaccharides (Lage and Bondoso 2014). Algal sulfated
573 polysaccharides were also found in some microalgae, including diatoms and dinoflagellates
574 (Raposo, De Morais and Bernardo De Morais 2013). *Gyrodinium*, a dinoflagellate that produces
575 algal sulfated polysaccharides (Raposo, De Morais and Bernardo De Morais 2013), was found to
576 be enriched (between 30 – 40%) in the upper oxycline of the ETNP OMZ in the small (1.6 – 30
577 μm) and large particle-associated (> 30 μm) size fraction (Duret *et al.* 2015). The high relative
578 abundance of OM190 group and Phycisphaerales in the ETNP where macroalgae abundance is
579 low suggests that these Planctomycetota lineages may feed on biopolymers synthesized by
580 microalgae, such as *Gyrodinium*. Brocadiales, which includes anammox bacteria (e.g. *Scalindua*
581 and *Brocadia*), were the only Planctomycetota lineage that was primarily free-living
582 (Supplementary Figure S1). While the vertical sampling resolution in our study did not cover the
583 core of the anoxic layer, the maxima in the relative abundance of Brocadiales near the oxic-
584 anoxic interface at Stations 1 and 2 corresponded to elevated anammox rates measured during
585 the same cruise (Tracey *et al.* 2022). On the other hand, the relative abundance of Brocadiales
586 was very low at the coastal Station 3, corresponding to lower anammox rates than the offshore
587 stations, potentially because the high surface primary production favored denitrification over
588 anammox (Figure 2) (Tracey *et al.* 2022).
589
590 In addition to Flavobacteriales, Verrucomicrobiota, and Planctomycetota, other lineages that
591 marked a distinct shift along the redoxcline included Desulfobacterota (formerly under
592 Deltaproteobacteria), Marinimicrobia (formerly known as SAR406), and Gammaproteobacteria

593 (*Aquabacterium*, *Curvibacter*, and *Woeseia*) (Figure 5). These lineages potentially mediate key
594 biogeochemical processes in the oxycline and the anoxic depths. For example, Desulfobacterota
595 may be mediating sulfate reduction, nitrogen fixation, and dissimilatory nitrate reduction to
596 ammonia (DNRA) (Langwig *et al.* 2022), Marinimicrobia may be mediating partial
597 denitrification (Bertagnolli *et al.* 2017; Hawley *et al.* 2017), and Gammaproteobacteria may be
598 mediating denitrification (Patureau *et al.* 1994; Zielińska *et al.* 2016; Mußmann *et al.* 2017; Xu
599 *et al.* 2022).
600
601 ***Insights from microbial co-occurrence network analyses***
602 Microbial co-occurrence networks have become a staple tool among microbial ecologists to infer
603 potential interactions between community members and aid in hypothesis generation (Banerjee,
604 Schlaeppi and Van Der Heijden 2018; Röttjers and Faust 2018). Many tools have been
605 developed and used broadly, such as SparCC (Friedman and Alm 2012), SpiecEasi (Kurtz *et al.*
606 2015), MENAP (Deng *et al.* 2012), and WGCNA (Langfelder and Horvath 2008). We selected
607 MENAP for our analyses because it is robust to noise and it has been updated with the iDIRECT
608 module that excels at detecting indirect interactions from networks (Xiao *et al.* 2022). The
609 removal of indirect interactions between prokaryotic community members resulted in three
610 networks with a smaller number of nodes and edges compared to many previously reported
611 planktonic networks (Milici *et al.* 2016; Beman *et al.* 2021; Guo *et al.* 2022; Wu *et al.* 2022).
612 Nevertheless, the microbial co-occurrence networks we reconstructed revealed previously
613 overlooked potential interactions between prokaryotic members that potentially impact key
614 biogeochemical processes in the ETNP OMZ.
615

616 The microbial communities from three size fractions formed networks distinct from each other,
617 with the largest number of nodes and the largest number of modules identified in the free-living
618 community and the largest number of potential interactions and smallest number of modules in
619 the large particle-associated community (Figure 6). This result is intuitive as the average physical
620 distance between any two free-living prokaryotic cells (464 μm in 1 ml of seawater containing
621 10^5 cells) is at least one order of magnitude higher than that between two prokaryotic cells
622 attached to large particles (< 22 μm), which translates to lower probability for two cells to
623 encounter in the free-living than the large particle-associated community. Most of the
624 correlations (57%) in the free-living community were positive, whereas a smaller percentage of
625 correlations in the small particle-associated (48%) and large particle-associated (46%)
626 communities were positive. We interpret this as greater competition for substrates among
627 prokaryotic community members on large and small particles than their free-living counterparts.
628
629 In all three networks, putative keystone species included SAR11 (both Clades I and II),
630 consistent with their role as the most abundant heterotrophic prokaryotes in the world's oceans
631 (Vaulot *et al.* 1995; Partensky, Hess and Vaulot 1999; Giovannoni 2017). Most of the SAR11
632 subclades (i.e. clade Ia, Ib, and II) are present in all depths sampled. However, SAR11 subclades
633 showed distinct depth distribution patterns in the majority of samples throughout the ETNP
634 OMZ, with clades Ia and II dominating in the surface, oxic environment, and clade Ib increasing
635 in relative abundance at the oxycline and anoxic depths (Supplementary Figure S11). The
636 distribution of these different SAR11 ecotypes is consistent with previous studies of the ETSP
637 (Wright, Konwar and Hallam 2012; Aldunate *et al.* 2018). *Prochlorococcus*, *Synechococcus*, and
638 chloroplast DNA (classified as uncultured marine eukaryotes) were identified as putative

639 keystone species in the free-living and small particle-associated networks, consistent with their
640 role as the most abundant photoautotrophic prokaryotes in the world's oceans (Vaulot *et al.*
641 1995; Partensky, Hess and Vaulot 1999; Giovannoni 2017). Both *Prochlorococcus* and
642 *Synechococcus* were a significant part of the prokaryotic communities at all size fractions, each
643 accounting for up to ~20% by relative abundance (Supplementary Figure S12). As the primary
644 photoautotroph in the ETNP OMZ, the high connectedness of Cyanobacteriotal ASVs
645 underscores their essential role in the food web as well as their contribution to oxygen
646 production in the DCM, which fuels microaerophilic processes, including nitrification and
647 denitrification (Garcia-Robledo *et al.* 2017).
648
649 In free-living communities, many positive correlations are between Cyanobacteriota (i.e.
650 *Synechococcus* and chloroplast DNA) and SAR11 clade I (Figure 6). SAR86 (identified as a
651 putative keystone species, ASV43) also had positive correlations with Cyanobacteriota (i.e.
652 *Prochlorococcus* and *Synechococcus*). SAR86 are prevalent in oxic waters above the ETNP
653 OMZ and positively correlated with oxygen concentration (Guo *et al.* 2022), similar to the
654 distribution of Cyanobacteriota (Figure 3). These positive correlations suggest that SAR86 and
655 SAR11 may have a shared niche with Cyanobacteriota or benefit from the carbon fixed and
656 released by Cyanobacteriota. On the other hand, negative correlations between chloroplast DNA
657 (potentially unclassified photosynthetic picoeukaryotes) and a few heterotrophic ASVs, such as
658 SAR324 and the AEGEAN-169 marine group (Rhodospirillales), suggest varying niches or
659 potential competition for nutrients such as fixed nitrogen used for denitrification or fermentation
660 of organic matter (Mulla *et al.* 2018; Bandekar et al., 2018; Lüke et al., 2016). These inferred

661 interactions between phytoplankton and heterotrophic bacteria require further experimentation to
662 be verified.
663
664 The small particle-associated (2 – 22 μm) network included seven modules with the fewest
665 identified nodes and connections. Two fungal ASVs (*Rhodotorula sp.* (Basidiomycota, ASV-F1)
666 and *Cladosporium sp.* (Ascomycota, ASV-F74) from the same filters (Peng and Valentine 2021)
667 were included in the small particle-associated network (Figure 7). The *Cladosporium sp.* ASV
668 (ASV-F74) was identified as a putative keystone species (Supplementary Table S10), with a
669 positive correlation with *Rhodotorula sp.* and a negative correlation with SAR11. *Cladosporium*
670 *sp.* is a cosmopolitan and diverse hyphomycete genus found commonly in all environments
671 (Bensch *et al.* 2012; Peng *et al.* 2021). *Cladosporium sp.* was found predominantly in the large
672 particle-associated size fraction in the ETNP OMZ (Peng and Valentine 2021). *Cladosporium*
673 can degrade a wide range of substrates including polymeric organic matter, polycyclic aromatic
674 hydrocarbons, and polyurethane (Birolli *et al.* 2018; Masigol *et al.* 2019; Zhang *et al.* 2022).
675 Marine *Cladosporium* has been found to produce secondary metabolites (Lee *et al.* 2023), which
676 can be deployed in interspecies competition. *Rhodotorula sp*. is prevalent and abundant in the
677 ETNP, specifically in the free-living size fraction (Peng and Valentine 2021) and Arabian Sea
678 OMZ (Fell 1967). The positive connection between *Rhodotorula sp.* and *Cladosporium sp.*
679 suggests that they have a shared niche in the ocean.
680
681 The large particle-associated (> 22 μm) network consisted of the largest number of edges
682 connecting them, but only three modules were identified by our network analysis, and the
683 average path distance was the lowest (Supplementary Table S8), suggesting either bacteria/fungi

684 being attached to the same particle or close interactions between microbes in large particle-
685 associated communities. In contrast to free-living and small particle-associated networks,
686 alphaproteobacterial ASVs were no longer the dominant taxa, as there were more ASVs from
687 Gammaproteobacteria and Bacteroidota included in the large particle-associated network than
688 Alphaproteobacterial ASVs (Figure 8). This shift in the network community composition is
689 consistent with the shift in the overall microbial community composition from free-living to
690 particle-associated size fractions discussed above (Figure 5). The two most abundant
691 Gammaproteobacterial taxa in the large particle-associated network were Alteromonadales and
692 Vibrionales (Figure 8), which included one putative keystone species (Supplementary Table
693 S10), ASV130 (*Vibrionaceae*). Alteromonadales are known to use phytoplankton-derived
694 polysaccharide microgels (Taylor & Cunliffe, 2017), and they contribute to carbon cycling in
695 both free-living and particle-associated fractions at suboxic ($[O_2] < 5$ μM) depths of OMZs
696 (Henríquez-Castillo *et al.* 2022). *Vibrionaceae* (Order: Vibrionales, Class:
697 Gammaproteobacteria) is ubiquitous in coastal and open oceans (Moi *et al.* 2017). We
698 hypothesize that in the ETNP OMZ, the putative keystone *Vibrionaceae* (ASV130,
699 Supplementary Table S10) contributes to nitrogen fixation, as isolated strains of Vibrionales are
700 known nitrogen fixers (Tibbles and Rawlings 1994; Cheung *et al.* 2016), and the nitrogenase
701 gene from *Vibrio* is prevalent and abundant in both the oligotrophic and coastal oceans (Zehr,
702 Mellon and Zani 1998). *Rhodotorula sp.* (ASV-F1) was positively connected with other
703 *Rhodotorula sp.* (ASV-F4) and Rhodospirillales (Alphaproteobacteria), suggesting these
704 organisms are growing on the same particle or shared niche between the same fungal genus and
705 some Alphaproteobacteria. *Rhodotorula sp.* negatively connected with SAR11, Vibrionales, and

706 SAR324, suggesting that marine fungi and these bacteria may compete for substrates on large
707 particles or have varying niches (Supplementary Figures S13 – 14).
708
709 **Conclusions**
710 In this study, large volumes (23 - 69 L, Supplementary Table S1) were filtered through three size
711 fractions (0.22 μm, 2.0 μm, and 22 μm) to separate both small and large particle-associated
712 microbes from free-living microbial communities from the eastern tropical North Pacific (ETNP)
713 OMZ. By sequencing and analyzing the V4 region of the 16S rRNA gene from all three size
714 fractions, we revealed previously overlooked lineages that drive the differences in community
715 composition between different size fractions and under different oxygen levels. ASVs from
716 Flavobacteriales, Planctomycetota, and Verrucomicrobiota were the key members that drove the
717 differences in microbial community composition between different oxygen regimes. As these
718 lineages have been reported as particle-associated microbes, we found Flavobacteriales,
719 Planctomycetota, and Verrucomicrobiota to have a higher relative abundance than previously
720 reported in the ETNP OMZ. By including previously published fungal community composition
721 data in network analyses, we identified both prokaryotic and fungal (*Cladosporium*
722 (Ascomycota)) putative keystone species that likely play an important role in OMZ
723 biogeochemistry. The network analyses at these three size fractions suggest that potential
724 interactions between fungi and prokaryotes in the ocean take place primarily on particles. Our
725 findings provide a rich resource for generating hypotheses regarding microbial interactions in
726 oceanic OMZs.

Acknowledgments

We are indebted to members of Bess Ward's and Karen Casciotti's Labs, the crew of R/V Sally Ride, and Frank Kinnaman for general assistance. This work is supported by the Simons Foundation Postdoctoral Fellowship in Marine Microbial Ecology (No. 547606) and the Simons Early Career Investigator in Aquatic Microbial Ecology and Evolution Award (No. SFI-LS-ECIAMEE-00006491) to Xuefeng Peng and NSF Grants OCE-1635562 and OCE-1756947, and the C-BRIDGES program to David Valentine. We would like to acknowledge that the Research Computing program under the Division of Information Technology at the University of South Carolina contributed to the results in this research by providing High Performance Computing resources and expertise.

References

Aldunate M, De La Iglesia R, Bertagnolli AD *et al.* Oxygen modulates bacterial community composition in the coastal upwelling waters off central Chile. *Deep Sea Research Part II: Topical Studies in Oceanography* 2018;**156**:68–79.

Bandekar M, Ramaiah N, Jain A *et al. Distinctly Different Bacterial Communities in Surface and Oxygen Minimum Layers in the Arabian Sea*. Biodiversity and Ecosystem Function: Marine, 2016.

Bandekar M, Ramaiah N, Jain A *et al.* Seasonal and depth-wise variations in bacterial and archaeal groups in the Arabian Sea oxygen minimum zone. *Deep Sea Research Part II: Topical Studies in Oceanography* 2018;**156**:4–18.

Bandekar M, Ramaiah N, Meena RM. Diversity and abundance of denitrifying and anammox bacteria from the Arabian Sea oxygen minimum zone. *Deep Sea Research Part II: Topical Studies in Oceanography* 2018;**156**:19–26.

Banerjee S, Schlaeppi K, Van Der Heijden MGA. Keystone taxa as drivers of microbiome structure and functioning. *Nat Rev Microbiol* 2018;**16**:567–76.

Beman JM, Carolan MT. Deoxygenation alters bacterial diversity and community composition in the ocean's largest oxygen minimum zone. *Nat Commun* 2013;**4**:2705.

Beman JM, Vargas SM, Vazquez S *et al.* Biogeochemistry and hydrography shape microbial community assembly and activity in the eastern tropical North Pacific Ocean oxygen minimum zone. *Environmental Microbiology* 2021;**23**:2765–81.

Bensch K, Braun U, Groenewald JZ *et al.* The genus Cladosporium. *Studies in Mycology* 2012;**72**:1–401.

Bertagnolli AD, Padilla CC, Glass JB *et al.* Metabolic potential and *in situ* activity of marine Marinimicrobia bacteria in an anoxic water column. *Environ Microbiol* 2017;**19**:4392–416.

Bertagnolli AD, Stewart FJ. Microbial niches in marine oxygen minimum zones. *Nat Rev Microbiol* 2018;**16**:723–9.

Birolli WG, de A. Santos D, Alvarenga N *et al.* Biodegradation of anthracene and several PAHs by the marine-derived fungus Cladosporium sp. CBMAI 1237. *Marine Pollution Bulletin* 2018;**129**:525–33.

Borcard D, Gillet F, Legendre P. *Numerical Ecology with R*. New York, NY: Springer New York, 2011.

Bowman JP. *Crocinitomix*. In: Whitman WB, Rainey F, Kämpfer P, et al. (eds.). *Bergey's Manual of Systematics of Archaea and Bacteria*. 1st ed. Wiley, 2015, 1–5.

774 Braman RS, Hendrix SA. Nanogram nitrite and nitrate determination in environmental and
775 biological materials by vanadium(III) reduction with chemiluminescence detection. *Anal*
776 *Chem* 1989;**61**:2715–8.

777 Breitburg D, Levin LA, Oschlies A *et al.* Declining oxygen in the global ocean and coastal
778 waters. *Science* 2018;**359**:eaam7240.

779 Bryant JA, Stewart FJ, Eppley JM *et al.* Microbial community phylogenetic and trait diversity
780 declines with depth in a marine oxygen minimum zone. *Ecology* 2012;**93**:1659–73.

781 Bueno E, Mesa S, Bedmar EJ *et al.* Bacterial Adaptation of Respiration from Oxic to Microoxic
782 and Anoxic Conditions: Redox Control. *Antioxidants & Redox Signaling* 2012;**16**:819–
783 52.

784 Callbeck CM, Ehrenfels B, Baumann KBL *et al.* Anoxic chlorophyll maximum enhances local
785 organic matter remineralization and nitrogen loss in Lake Tanganyika. *Nat Commun*
786 2021;**12**:830.

787 Cameron ES, Schmidt PJ, Tremblay BJ-M *et al.* Enhancing diversity analysis by repeatedly
788 rarefying next generation sequencing data describing microbial communities. *Sci Rep*
789 2021;**11**:22302.

790 Canfield DE, Kraft B. The 'oxygen' in oxygen minimum zones. *Environmental Microbiology*
791 2022;**24**:5332–44.

792 Caporaso JG, Ackermann G, Apprill A *et al. EMP 16S Illumina Amplicon Protocol V1*., 2018.

793 Caporaso JG, Kuczynski J, Stombaugh J *et al.* QIIME allows analysis of high-throughput
794 community sequencing data. *Nat Methods* 2010;**7**:335–6.

795 Caporaso JG, Lauber CL, Walters WA *et al.* Ultra-high-throughput microbial community
796 analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal* 2012;**6**:1621–4.

797 Carolan MT, Smith JM, Beman JM. Transcriptomic evidence for microbial sulfur cycling in the
798 eastern tropical North Pacific oxygen minimum zone. *Front Microbiol* 2015;**06**, DOI:
799 10.3389/fmicb.2015.00334.

800 Chavda ND, Deota HN, Kota VKB. Poisson to GOE transition in the distribution of the ratio of
801 consecutive level spacings. *Physics Letters A* 2014;**378**:3012–7.

802 Chen I-MA, Chu K, Palaniappan K *et al.* The IMG/M data management and analysis system v.7:
803 content updates and new features. *Nucleic Acids Research* 2023;**51**:D723–32.

804 Cheung S, Xia X, Guo C *et al.* Diazotroph community structure in the deep oxygen minimum
805 zone of the Costa Rica Dome. *J Plankton Res* 2016;**38**:380–91.

806 Codispoti LA, Friederich GE, Packard TT *et al.* High Nitrite Levels off Northern Peru: A Signal
807 of Instability in the Marine Denitrification Rate. *Science* 1986;**233**:1200–2.

808 Crenn K, Duffieux D, Jeanthon C. Bacterial Epibiotic Communities of Ubiquitous and Abundant
809 Marine Diatoms Are Distinct in Short- and Long-Term Associations. *Frontiers in*
810 *Microbiology* 2018;**9**.

811 Davies EJ, McKee D, Bowers D *et al.* Optically significant particle sizes in seawater. *Appl Opt*
812 2014;**53**:1067.

813 De Boyer Montégut C. Mixed layer depth over the global ocean: An examination of profile data
814 and a profile-based climatology. *J Geophys Res* 2004;**109**:C12003.

815 Deng Y, Jiang Y-H, Yang Y *et al.* Molecular ecological network analyses. *BMC Bioinformatics*
816 2012;**13**:113.

817 Duret MT, Pachiadaki MG, Stewart FJ *et al.* Size-fractionated diversity of eukaryotic microbial
818 communities in the Eastern Tropical North Pacific oxygen minimum zone. *FEMS*
819 *Microbiology Ecology* 2015;**91**, DOI: 10.1093/femsec/fiv037.

820 Faull LM, Mara P, Taylor GT *et al.* Imprint of Trace Dissolved Oxygen on Prokaryoplankton
821 Community Structure in an Oxygen Minimum Zone. *Front Mar Sci* 2020;**7**:360.

822 Fell J. Distribution of Yeasts in the Indian Ocean. *Bulletin of Marine Science* 1967;**17**:454–70.

823 Fernandes GL, Shenoy BD, Damare SR. Diversity of Bacterial Community in the Oxygen
824 Minimum Zones of Arabian Sea and Bay of Bengal as Deduced by Illumina Sequencing.
825 *Front Microbiol* 2020;**10**:3153.

826 Friedman J, Alm EJ. Inferring Correlation Networks from Genomic Survey Data. Von Mering C
827 (ed.). *PLoS Comput Biol* 2012;**8**:e1002687.

828 Fuchsman CA, Palevsky HI, Widner B *et al.* Cyanobacteria and cyanophage contributions to
829 carbon and nitrogen cycling in an oligotrophic oxygen-deficient zone. *ISME J*
830 2019;**13**:2714–26.

831 Galán A, Molina V, Thamdrup B *et al.* Anammox bacteria and the anaerobic oxidation of
832 ammonium in the oxygen minimum zone off northern Chile. *Deep Sea Research Part II:*
833 *Topical Studies in Oceanography* 2009;**56**:1021–31.

834 Ganesh S, Bristow LA, Larsen M *et al.* Size-fraction partitioning of community gene
835 transcription and nitrogen metabolism in a marine oxygen minimum zone. *ISME J*
836 2015;**9**:2682–96.

837 Ganesh S, Parris DJ, DeLong EF *et al.* Metagenomic analysis of size-fractionated picoplankton
838 in a marine oxygen minimum zone. *ISME J* 2014;**8**:187–211.

839 Garcia-Robledo E, Padilla CC, Aldunate M *et al.* Cryptic oxygen cycling in anoxic marine
840 zones. *Proc Natl Acad Sci USA* 2017;**114**:8319–24.

841 Giovannoni SJ. SAR11 Bacteria: The Most Abundant Plankton in the Oceans. *Annu Rev Mar Sci*
842 2017;**9**:231–55.

843 Guimerà R, Amaral LAN. Cartography of complex networks: modules and universal roles. *J Stat*
844 *Mech* 2005;**2005**:P02001.

845 Guo R, Ma X, Zhang J *et al.* Microbial community structures and important taxa across oxygen
846 gradients in the Andaman Sea and eastern Bay of Bengal epipelagic waters. *Front*
847 *Microbiol* 2022;**13**:1041521.

848 Hawley AK, Nobu MK, Wright JJ *et al.* Diverse Marinimicrobia bacteria may mediate coupled
849 biogeochemical cycles along eco-thermodynamic gradients. *Nat Commun* 2017;**8**:1507.

850 Henríquez-Castillo C, Plominsky AM, Ramírez-Flandes S *et al.* Metaomics unveils the
851 contribution of Alteromonas bacteria to carbon cycling in marine oxygen minimum
852 zones. *Front Mar Sci* 2022;**9**:993667.

853 Holmes RM, Aminot A, Kérouel R *et al.* A simple and precise method for measuring ammonium
854 in marine and freshwater ecosystems. *Can J Fish Aquat Sci* 1999;**56**:1801–8.

855 James CC, Barton AD, Allen LZ *et al.* Influence of nutrient supply on plankton microbiome
856 biodiversity and distribution in a coastal upwelling region. *Nat Commun* 2022;**13**:2448.

857 Kalvelage T, Lavik G, Jensen MM *et al.* Aerobic Microbial Respiration In Oceanic Oxygen
858 Minimum Zones. Quan Z-X (ed.). *PLoS ONE* 2015;**10**:e0133526.

859 Karstensen J, Stramma L, Visbeck M. Oxygen minimum zones in the eastern tropical Atlantic
860 and Pacific oceans. *Progress in Oceanography* 2008;**77**:331–50.

861 Kelly CL, Travis NM, Baya PA *et al.* Quantifying Nitrous Oxide Cycling Regimes in the Eastern
862 Tropical North Pacific Ocean With Isotopomer Analysis. *Global Biogeochem Cycles*
863 2021;**35**, DOI: 10.1029/2020GB006637.

864 Klawonn I, Bonaglia S, Brüchert V *et al.* Aerobic and anaerobic nitrogen transformation
865 processes in N2-fixing cyanobacterial aggregates. *ISME J* 2015;**9**:1456–66.

866 Kozich JJ, Westcott SL, Baxter NT *et al.* Development of a Dual-Index Sequencing Strategy and
867 Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina
868 Sequencing Platform. *Appl Environ Microbiol* 2013;**79**:5112–20.

869 Kurtz ZD, Müller CL, Miraldi ER *et al.* Sparse and Compositionally Robust Inference of
870 Microbial Ecological Networks. Von Mering C (ed.). *PLoS Comput Biol*
871 2015;**11**:e1004226.

872 Lage OM, Bondoso J. Planctomycetes and macroalgae, a striking association. *Front Microbiol*
873 2014;**5**, DOI: 10.3389/fmicb.2014.00267.

874 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis.
875 *BMC Bioinformatics* 2008;**9**:559.

876 Langwig MV, De Anda V, Dombrowski N *et al.* Large-scale protein level comparison of
877 Deltaproteobacteria reveals cohesive metabolic groups. *ISME J* 2022;**16**:307–20.

878 Lavin P, González B, Santibáñez JF *et al.* Novel lineages of Prochlorococcus thrive within the
879 oxygen minimum zone of the eastern tropical South Pacific: Picocyanobacteria in oxygen
880 minimum zones. *Environmental Microbiology Reports* 2010;**2**:728–38.

881 Lee W, Kim JS, Seo CW *et al.* Diversity of Cladosporium (Cladosporiales, Cladosporiaceae)
882 species in marine environments and report on five new species. *MC* 2023;**98**:87–111.

883 Legendre P, Borcard D. Box-Cox-chord transformations for community composition data prior
884 to beta diversity analysis. *Ecography* 2018;**41**:1820–4.

885 Long AM, Hou S, Ignacio-Espinoza JC *et al.* Benchmarking microbial growth rate predictions
886 from metagenomes. *The ISME Journal* 2021;**15**:183–95.

887 Maerz J, Six KD, Stemmler I *et al.* Microstructure and composition of marine aggregates as co-
888 determinants for vertical particulate organic carbon transfer in the global ocean.
889 *Biogeosciences* 2020;**17**:1765–803.

890 Masigol H, Khodaparast SA, Woodhouse JN *et al.* The contrasting roles of aquatic fungi and
891 oomycetes in the degradation and transformation of polymeric organic matter. *Limnology*
892 *and Oceanography* 2019;**64**:2662–78.

893 McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and
894 Graphics of Microbiome Census Data. Watson M (ed.). *PLoS ONE* 2013;**8**:e61217.

895 Mestre M, Ferrera I, Borrull E *et al.* Spatial variability of marine bacterial and archaeal
896 communities along the particulate matter continuum. *Molecular Ecology* 2017;**26**:6827–
897 40.

898 Milici M, Deng Z-L, Tomasch J *et al.* Co-occurrence Analysis of Microbial Taxa in the Atlantic
899 Ocean Reveals High Connectivity in the Free-Living Bacterioplankton. *Front Microbiol*
900 2016;**7**, DOI: 10.3389/fmicb.2016.00649.

901 Moi IM, Roslan NN, Leow ATC *et al.* The biology and the importance of Photobacterium
902 species. *Appl Microbiol Biotechnol* 2017;**101**:4371–85.

903 Molina V, Ulloa O, Farías L *et al.* Ammonia-Oxidizing β-Proteobacteria from the Oxygen
904 Minimum Zone off Northern Chile. *Appl Environ Microbiol* 2007;**73**:3547–55.

905 Mukherjee S, Stamatis D, Li CT *et al.* Twenty-five years of Genomes OnLine Database
906 (GOLD): data updates and new features in v.9. *Nucleic Acids Research* 2023;**51**:D957–
907 63.

908 Mußmann M, Pjevac P, Krüger K *et al.* Genomic repertoire of the Woeseiaceae/JTB255,
909 cosmopolitan and abundant core members of microbial communities in marine
910 sediments. *ISME J* 2017;**11**:1276–81.

911 Oksanen J, Simpson G, Blanchet FG *et al. Vegan Community Ecology Package Version 2.6-2*
912 *April 2022*., 2022.

913 Orellana LH, Francis TB, Ferraro M *et al.* Verrucomicrobiota are specialist consumers of
914 sulfated methyl pentoses during diatom blooms. *ISME J* 2022;**16**:630–41.

915 Orsi WD, Smith JM, Liu S *et al.* Diverse, uncultivated bacteria and archaea underlying the
916 cycling of dissolved protein in the ocean. *ISME J* 2016;**10**:2158–73.

917 Paerl HW, Otten TG. Harmful Cyanobacterial Blooms: Causes, Consequences, and Controls.
918 *Microbial Ecology* 2013;**65**:995–1010.

919 Pajares S. Unraveling the distribution patterns of bacterioplankton in a mesoscale cyclonic eddy
920 confined to an oxygen-depleted basin. *Aquat Microb Ecol* 2021;**87**:151–66.

921 Pajares S, Varona-Cordero F, Hernández-Becerril DU. Spatial Distribution Patterns of
922 Bacterioplankton in the Oxygen Minimum Zone of the Tropical Mexican Pacific. *Microb*
923 *Ecol* 2020;**80**:519–36.

924 Partensky F, Hess WR, Vaulot D. *Prochlorococcus* , a Marine Photosynthetic Prokaryote of
925 Global Significance. *Microbiol Mol Biol Rev* 1999;**63**:106–27.

926 Patureau D, Davison J, Bernet N *et al.* Denitrification under various aeration conditions in
927 Comamonas sp., strain SGLY2. *FEMS Microbiology Ecology* 1994;**14**:71–8.

928 Peng X, Fuchsman CA, Jayakumar A *et al.* Ammonia and nitrite oxidation in the Eastern
929 Tropical North Pacific. *Global Biogeochem Cycles* 2015;**29**:2034–49.

930 Peng X, Fuchsman CA, Jayakumar A *et al.* Revisiting nitrification in the Eastern Tropical S outh
931 P acific: A focus on controls. *J Geophys Res Oceans* 2016;**121**:1667–84.

932 Peng X, Gat D, Paytan A *et al.* The Response of Airborne Mycobiome to Dust Storms in the
933 Eastern Mediterranean. *JoF* 2021;**7**:802.

934 Peng X, Jayakumar A, Ward BB. Community composition of ammonia-oxidizing archaea from
935 surface and anoxic depths of oceanic oxygen minimum zones. *Front Microbiol* 2013;**4**,
936 DOI: 10.3389/fmicb.2013.00177.

937 Peng X, Valentine DL. Diversity and N2O Production Potential of Fungi in an Oceanic Oxygen
938 Minimum Zone. *JoF* 2021;**7**:218.

939 Peng X, Yousavich DJ, Bourbonnais A *et al.* The fate of fixed nitrogen in Santa Barbara Basin
940 sediments during seasonal anoxia. *Biogeosciences* 2024;**21**:3041–52.

941 Quast C, Pruesse E, Yilmaz P *et al.* The SILVA ribosomal RNA gene database project: improved
942 data processing and web-based tools. *Nucleic Acids Research* 2012;**41**:D590–6.

943 Rajpathak SN, Banerjee R, Mishra PG *et al.* An exploration of microbial and associated
944 functional diversity in the OMZ and non-OMZ areas in the Bay of Bengal. *J Biosci*
945 2018;**43**:635–48.

946 Rangamaran VR, Sankara Subramanian SH, Balachandran KRS *et al.* Vertical Microbial
947 Profiling of Arabian Sea Oxygen Minimal Zone Reveals Complex Bacterial
948 Communities and Distinct Functional Implications. *Microb Ecol* 2023;**85**:357–71.

949 Raposo M, De Morais R, Bernardo De Morais A. Bioactivity and Applications of Sulphated
950 Polysaccharides from Marine Microalgae. *Marine Drugs* 2013;**11**:233–52.

951 Read DA, Pietersen G. PCR bias associated with conserved primer binding sites, used to
952 determine genotype diversity within Citrus tristeza virus populations. *Journal of*
953 *Virological Methods* 2016;**237**:107–13.

954 Röttjers L, Faust K. From hairballs to hypotheses–biological insights from microbial networks.
955 *FEMS Microbiology Reviews* 2018;**42**:761–80.

956 Shannon P, Markiel A, Ozier O *et al.* Cytoscape: A Software Environment for Integrated Models
957 of Biomolecular Interaction Networks. *Genome Res* 2003;**13**:2498–504.

958 Silverman JD, Bloom RJ, Jiang S *et al.* Measuring and mitigating PCR bias in microbiota
959 datasets. McHardy AC (ed.). *PLoS Comput Biol* 2021;**17**:e1009113.

960 Stevens H, Ulloa O. Bacterial diversity in the oxygen minimum zone of the eastern tropical
961 South Pacific. *Environ Microbiol* 2008;**10**:1244–59.

962 Stramma L, Johnson GC, Sprintall J *et al.* Expanding Oxygen-Minimum Zones in the Tropical
963 Oceans. *Science* 2008;**320**:655–8.

964 Strickland J, Parsons T. A Practical Handbook of Seawater Analysis. Ottawa. *Int Revue ges*
965 *Hydrobiol Hydrogr* 1970;**55**:167–167.

966 Suter EA, Pachiadaki M, Taylor GT *et al.* Free-living chemoautotrophic and particle-attached
967 heterotrophic prokaryotes dominate microbial assemblages along a pelagic redox
968 gradient. *Environ Microbiol* 2018;**20**:693–712.

969 Taylor BW, Keep CF, Hall RO *et al.* Improving the fluorometric ammonium method: matrix
970 effects, background fluorescence, and standard additions. *Journal of the North American*
971 *Benthological Society* 2007;**26**:167–77.

972 Taylor JD, Cunliffe M. Coastal bacterioplankton community response to diatom-derived
973 polysaccharide microgels: Polysaccharide microgels and coastal bacterioplankton.
974 *Environmental Microbiology Reports* 2017;**9**:151–7.

975 Teeling H, Fuchs BM, Bennke CM *et al.* Recurring patterns in bacterioplankton dynamics during
976 coastal spring algae blooms. *eLife* 2016;**5**:e11888.

977 Tibbles BJ, Rawlings DE. Characterization of nitrogen-fixing bacteria from a temperate
978 saltmarsh lagoon, including isolates that produce ethane from acetylene. *Microb Ecol*
979 1994;**27**, DOI: 10.1007/BF00170115.

980 Torres-Beltrán M, Mueller A, Scofield M *et al.* Sampling and Processing Methods Impact
981 Microbial Community Structure and Potential Activity in a Seasonally Anoxic Fjord:
982 Saanich Inlet, British Columbia. *Front Mar Sci* 2019;**6**:132.

983 Tracey JC, Babbin AR, Wallace E *et al. All about Nitrite: Exploring Nitrite Sources and Sinks in*
984 *the Eastern Tropical North Pacific Oxygen Minimum Zone*. Biogeochemistry: Open
985 Ocean, 2022.

986 Travis NM, Kelly CL, Mulholland MR *et al.* Nitrite cycling in the primary nitrite maxima of the
987 eastern tropical North Pacific. *Biogeosciences* 2023;**20**:325–47.

988 Tully BJ, Graham ED, Heidelberg JF. The reconstruction of 2,631 draft metagenome-assembled
989 genomes from the global oceans. *Sci Data* 2018;**5**:170203.

990 Turner JT. Zooplankton fecal pellets, marine snow and sinking phytoplankton blooms. *Aquatic*
991 *Microbial Ecology* 2002;**27**:57–102.

992 Ulloa O, Canfield DE, DeLong EF *et al.* Microbial oceanography of anoxic oxygen minimum
993 zones. *Proc Natl Acad Sci USA* 2012;**109**:15996–6003.

994 Unfried F, Becker S, Robb CS *et al.* Adaptive mechanisms that provide competitive advantages
995 to marine bacteroidetes during microalgal blooms. *ISME J* 2018;**12**:2894–906.

996 Vaulot D, Marie D, Olson RJ *et al.* Growth of *Prochlorococcus* , a Photosynthetic Prokaryote, in
997 the Equatorial Pacific Ocean. *Science* 1995;**268**:1480–2.

998 Vuillemin A, Coskun ÖK, Orsi WD. Microbial Activities and Selection from Surface Ocean to
999 Subseafloor on the Namibian Continental Shelf. Glass JB (ed.). *Appl Environ Microbiol*
1000 2022;**88**:e00216-22.

1001 Wright JJ, Konwar KM, Hallam SJ. Microbial ecology of expanding oxygen minimum zones.
1002 *Nat Rev Microbiol* 2012;**10**:381–94.

1003 Wu C, Narale DD, Cui Z *et al.* Diversity, structure, and distribution of bacterioplankton and
1004 diazotroph communities in the Bay of Bengal during the winter monsoon. *Front*
1005 *Microbiol* 2022;**13**:987462.

1006 Xiao N, Zhou A, Kempher ML *et al.* Disentangling direct from indirect relationships in
1007 association networks. *Proc Natl Acad Sci USA* 2022;**119**:e2109995119.

1008 Xu L, Su J, Ali A *et al.* Magnetite-loaded rice husk biochar promoted the denitrification
1009 performance of Aquabacterium sp. XL4 under low carbon to nitrogen ratio: Optimization
1010 and mechanism. *Bioresource Technology* 2022;**348**:126802.

1011 Zehr JP, Mellon MT, Zani S. New Nitrogen-Fixing Microorganisms Detected in Oligotrophic
1012 Oceans by Amplification of Nitrogenase (nifH) Genes. *Applied and Environmental*
1013 *Microbiology* 1998;**64**:3444–50.

1014 Zhang K, Hu J, Yang S *et al.* Biodegradation of polyester polyurethane by the marine fungus
1015 Cladosporium halotolerans 6UPA1. *Journal of Hazardous Materials* 2022;**437**:129406.

1016 Zhou J, Deng Y, Luo F *et al.* Functional Molecular Ecological Networks. Relman DA (ed.).
1017 *mBio* 2010;**1**:e00169-10.

1018 Zhou J, Deng Y, Luo F *et al.* Phylogenetic Molecular Ecological Network of Soil Microbial
1019 Communities in Response to Elevated CO_2. Relman D (ed.). *mBio* 2011;**2**:e00122-11.

1020 Zielińska M, Rusanowska P, Jarząbek J *et al.* Community dynamics of denitrifying bacteria in
1021 full-scale wastewater treatment plants. *Environmental Technology* 2016;**37**:2358–67.

1022

1023 **Data availability**

1024 The code, including a metadata sample, is available in the following repository:

1025 https://github.com/mthompson19/ETNP2018_16s. Raw reads generated in this study are

1026 available at the National Center for Biotechnology Information (NCBI) under BioProject

1027 PRJNA911303.

1028

1029 **Conflict of interest:** The authors declare no conflict of interest.

1030
1031

Figure captions For Figures 1-8

Figure 1: Location of Station 1 (the peripheral OMZ, 247°E, 10°N), Station 2 (the offshore OMZ, 255°E, 15.76°N), and Station 3 (the coastal OMZ, 257.65°E, 17.68°N), and oxygen concentrations at 150 m in the Eastern tropical north pacific (ETNP) oxygen minimum zone (OMZ).

Figure 2: Oxygen Concentrations, Chlorophyll a, and sigma-theta at Station 1 (the peripheral OMZ; A), Station 2 (the offshore OMZ; B), and Station 3 (the coastal OMZ; C) in the eastern tropical North Pacific (ETNP) oxygen minimum zone (OMZ). Nitrite (NO_2^-), nitrate (NO_3^-), and ammonium (NH_4^+) at the sampling depths at Station 1 (the peripheral OMZ, D), Station 2 (the offshore OMZ), E), and Station 3 (the coastal OMZ, F). NH_4^+ and NO_2^- concentration profiles were adapted from Travis et al. (2022). Red diamonds represent the sampling depth at each station. The dashed line indicates the mixed layer depth.

Figure 3: The relative abundance of the 1100 most abundant prokaryotic amplicon sequence variant (ASV) in A) Station 1 free-living, B) Station 1 small particle-associated, C) Station 1 filter large particle-associated, D) Station 2 free-living, E) Station 2 small particle-associated, F) Station 2 large particle-associated, G) Station 3 free-living, H) Station 3 small particle-associated, I) Station 3 large particle-associated in the ETNP OMZ. The large particle-associated sample from 70 m at Station 3 (marked as "NA") yielded < 20 raw sequencing reads and was excluded from the analysis.

1055 **Figure 4**: Distance triplot of redundancy analysis (RDA) on prokaryotic community composition
1056 in free-living, small particle-associated, and large particle-associated communities from the
1057 ETNP OMZ, using dissolved oxygen (O_2), ammonium concentration (NH_4^+), nitrite
1058 concentration (NO_2^-), temperature, and chlorophyll a concentration (approximated by
1059 fluorescence measured by a Seapoint chlorophyll fluorometer) as explanatory variables.
1060 Concentrations of each physicochemical variable at each depth are described in Supplementary
1061 Table S11. The blue arrows are the vectors of the explanatory variables. Circles represent
1062 samples from the coastal OMZ, triangles represent samples from the offshore OMZ, and squares
1063 represent samples from the peripheral OMZ. Each symbol was color shaded by the different
1064 oxygen regimes (surface/oxic (O_2 >150 μM), upper oxycline (detection limit < O_2 < 150 μM),
1065 anoxic (O_2 < detection limit), and lower oxycline (O_2 < detection limit and below anoxic
1066 depths)). The size of each symbol is dependent on the size fraction.
1067
1068 **Figure 5**. The normalized (by trimmed mean of M values) abundance (counts per million
1069 calculated by EdgeR) of 46 most abundant amplicon sequence variants (rows) that were
1070 differentially abundant between surface, upper oxycline and anoxic depths, and lower oxycline
1071 (L.O.) depths in all samples (columns) collected from the ETNP OMZ in the free-living (FL),
1072 small particle-associated (SPA), and large particle-associated (LPA) size fractions.
1073
1074 **Figure 6:** Co-occurrence network interactions of free-living microbes in the ETNP OMZ. Each
1075 node is labeled with an ASV code. The area of each node, which represents an amplicon
1076 sequence variant (ASV), is proportional to the non-transformed relative abundance of that ASV
1077 (the smallest node (ASV2049) has a relative abundance of 0.0000309%, and the largest node
1078 (ASV2) has a relative abundance of 0.0233%). Positive interactions are shown in red and

negative ones in blue. A teal outline around the symbol highlights ASVs identified as putative keystone species ($Z_i > 1.5$ and $P_i < 0.6$).

Figure 7: Co-occurrence network interactions of small particle-associated microbes in the ETNP OMZ. Each node is labeled with an ASV code. The area of each node, which represents a prokaryotic or fungal amplicon sequence variant (ASV), is proportional to the non-transformed relative abundance of that ASV (the smallest node (ASV370) has a relative abundance of 0.000348%, and the largest node (ASV2) has a relative abundance of 0.023%). Positive interactions are shown in red and negative ones in blue. A teal outline around the symbol highlights ASVs identified as putative network hubs ($Z_i > 1.5$ and $P_i < 0.6$).

Figure 8: Co-occurrence network interactions of large particle-associated microbes in the ETNP OMZ. Each node is labeled with an ASV code. The area of each node, which represents a prokaryotic or fungal amplicon sequence variant (ASV), is proportional to the non-transformed relative abundance of that ASV (the smallest node (ASV1040) has a relative abundance of 0.0000956%, and the largest node (ASV4) has a relative abundance of 0.0148%). Positive interactions are shown in red and negative ones in blue. A teal outline around the symbol highlights ASVs identified as putative network hubs ($Z_i > 1.5$ and $P_i < 0.6$).

1099 **Table 1:** Filtration volumes, size fractions, and their respective study from the three main OMZs
1100 (ETNP, ETSP, and the Arabian Sea).

Site	Filtration volume (L)	Size fractions used	Reference
Eastern tropical North Pacific (ETNP)	**23 - 69**	**0.22 µm, 2 µm, and 22 µm**	**This study**
	23 - 55	0.22 µm, 2 µm, and 22 µm	Peng & Valentine, 2021
	10	0.2 µm	Beman & Carolan, 2013
	2	0.2 µm	Beman et al., 2021
	2.5	0.2 µm	Carolan et al., 2015
	2 – 5	0.2 µm and 2.7 µm	Faull et al., 2020
	10 – 15	0.2 µm and 1.6 µm	Ganesh et al., 2015
	1.2	0.2 µm	Pajares et al., 2020
Eastern tropical South Pacific (ETSP)	10	0.2 µm and 3 µm	Galán et al., 2009
	10	0.2 µm, 5 µm, and 20 µm	Molina et al., 2007
	10	0.2 µm and 1.6 µm	Ganesh et al., 2014
	10	0.2 µm and 3 µm	Stevens & Ulloa, 2008
Arabian Sea	2.5	0.2 µm	Bandekar et al., 2016
	2.5	0.2 µm	Bandekar et al., 2018
	5	0.2 µm	Fernandes et al. 2019
	10	0.2 µm	Fernandes et al., 2020
	2	0.2 µm	Rajpathak et al., 2018
	2.5	0.2 µm	Rangamaran et al., 2023

1101

1102

1103 **Table 2**: Free-living, small particle-associated, and large particle-associated networks and their
1104 respective number of nodes (ASVs), edges (links), and positive edges (%).

Co-occurrence network size fraction	Number of ASVs (nodes)	Number of edges	Positive edges (%)
Free-living	132	232	56.5
Small particle-associated	51	79	48.1
Large particle-associated	68	687	46.0

1105

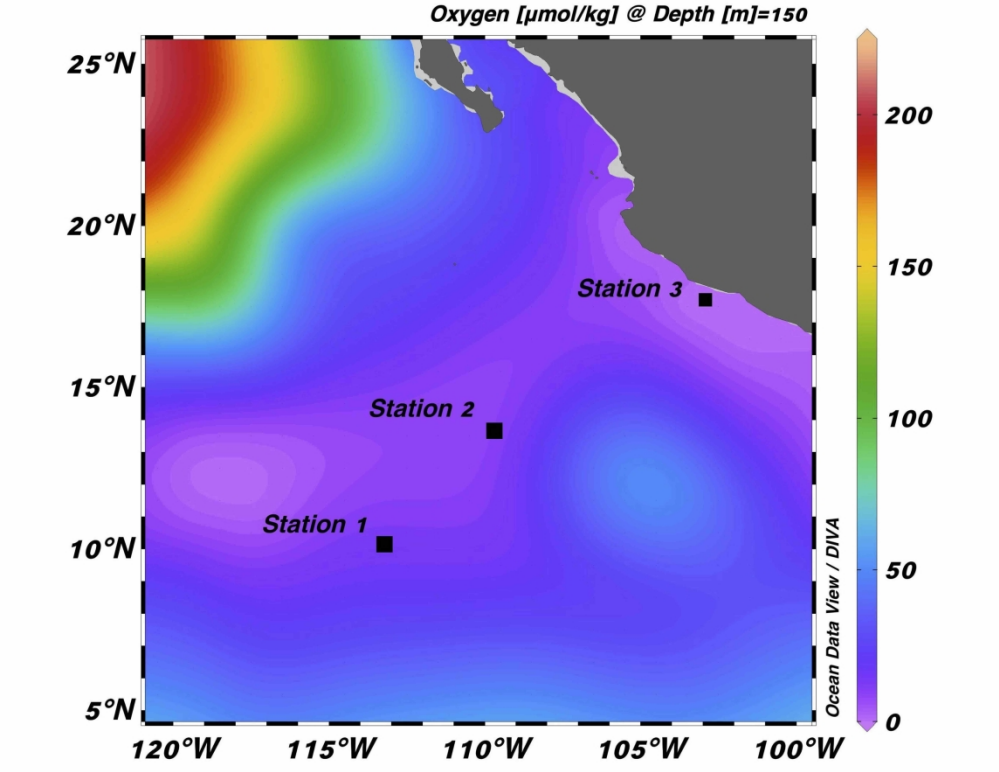

Figure 1: Location of Station 1 (the peripheral OMZ, 247°E, 10°N), Station 2 (the offshore OMZ, 255°E, 15.76°N), and Station 3 (the coastal OMZ, 257.65°E, 17.68°N), and oxygen concentrations at 150 m in the Eastern tropical north pacific (ETNP) oxygen minimum zone (OMZ).

279x215mm (300 x 300 DPI)

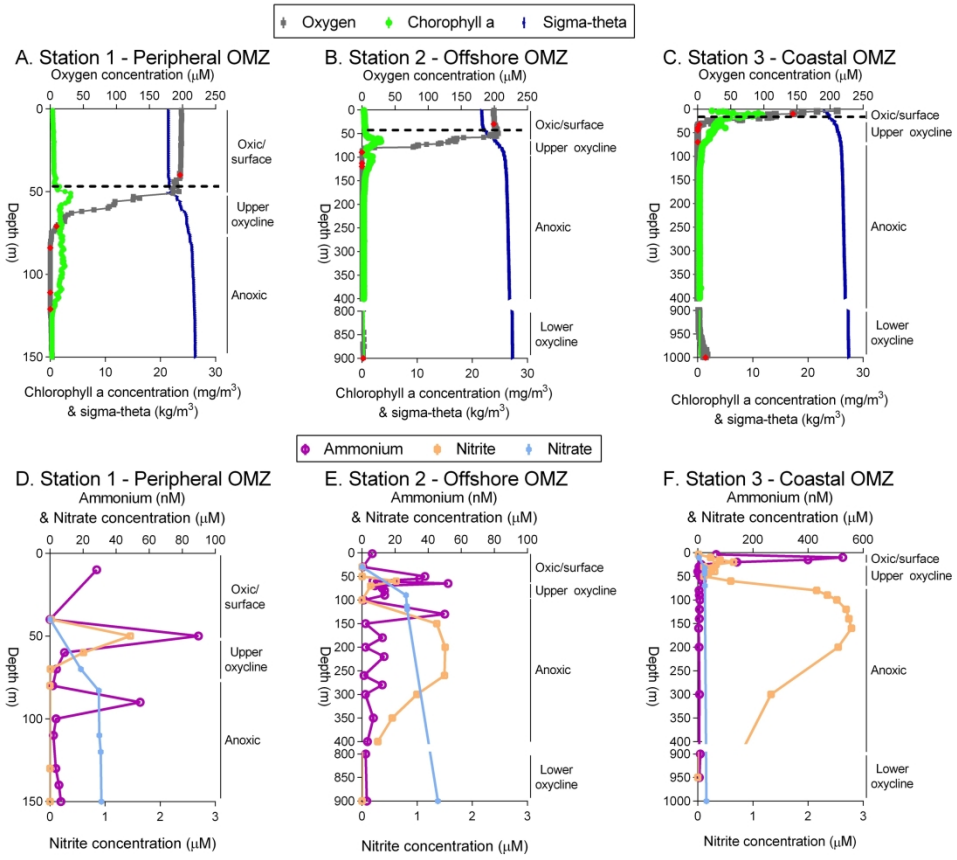

Figure 2: Oxygen Concentrations, Chlorophyll a, and sigma-theta at Station 1 (the peripheral OMZ; A), Station 2 (the offshore OMZ; B), and Station 3 (the coastal OMZ; C) in the eastern tropical North Pacific (ETNP) oxygen minimum zone (OMZ). Nitrite (NO2-), nitrate (NO3-), and ammonium (NH4+) at the sampling depths at Station 1 (the peripheral OMZ, D), Station 2 (the offshore OMZ), E), and Station 3 (the coastal OMZ, F). NH4+ and NO2- concentration profiles were adapted from Travis et al. (2022). Red diamonds represent the sampling depth at each station. The dashed line indicates the mixed layer depth.

245x219mm (300 x 300 DPI)

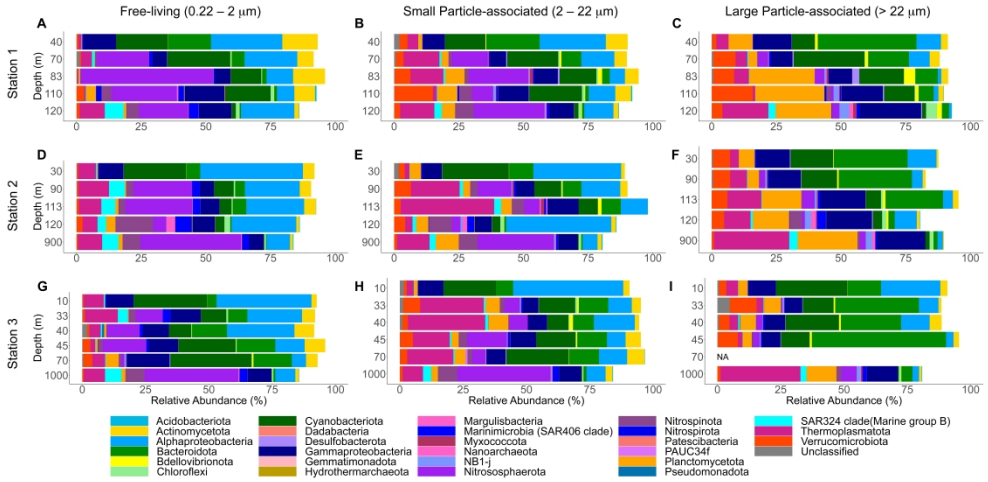

Figure 3: The relative abundance of the 1100 most abundant prokaryotic amplicon sequence variant (ASV) in A) Station 1 free-living, B) Station 1 small particle-associated, C) Station 1 filter large particle-associated, D) Station 2 free-living, E) Station 2 small particle-associated, F) Station 2 large particle-associated, G) Station 3 free-living, H) Station 3 small particle-associated, I) Station 3 large particle-associated in the ETNP OMZ. The large particle-associated sample from 70 m at Station 3 (marked as "NA") yielded < 20 raw sequencing reads and was excluded from the analysis.

1311x650mm (600 x 600 DPI)

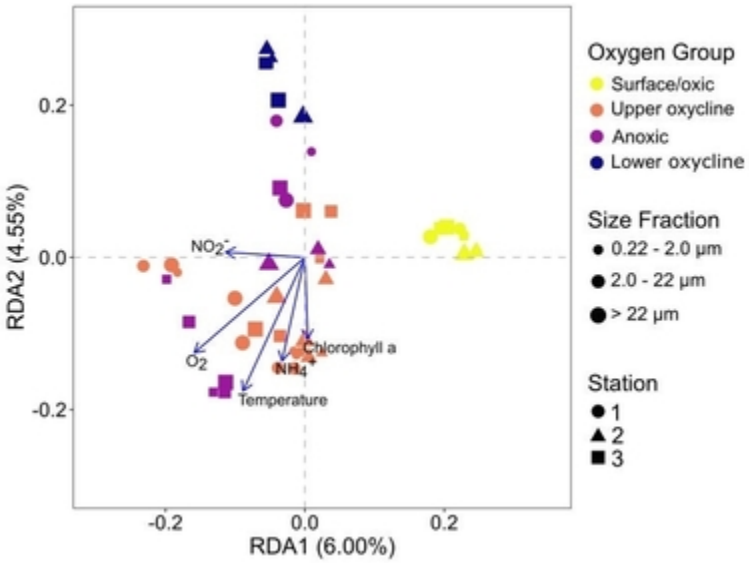

Figure 4: Distance triplot of redundancy analysis (RDA) on prokaryotic community composition in free-living, small particle-associated, and large particle-associated communities from the ETNP OMZ, using dissolved oxygen (O2), ammonium concentration (NH4+), nitrite concentration (NO2-), temperature, and chlorophyll a concentration (approximated by fluorescence measured by a Seapoint chlorophyll fluorometer) as explanatory variables. Concentrations of each physicochemical variable at each depth are described in Supplementary Table S11. The blue arrows are the vectors of the explanatory variables. Circles represent samples from the coastal OMZ, triangles represent samples from the offshore OMZ, and squares represent samples from the peripheral OMZ. Each symbol was color shaded by the different oxygen regimes (surface/oxic (O2 >150 µM), upper oxycline (detection limit < O2 < 150 µM), anoxic (O2 < detection limit), and lower oxycline (O2 < detection limit and below anoxic depths)). The size of each symbol is dependent on the size fraction.

32x23mm (300 x 300 DPI)

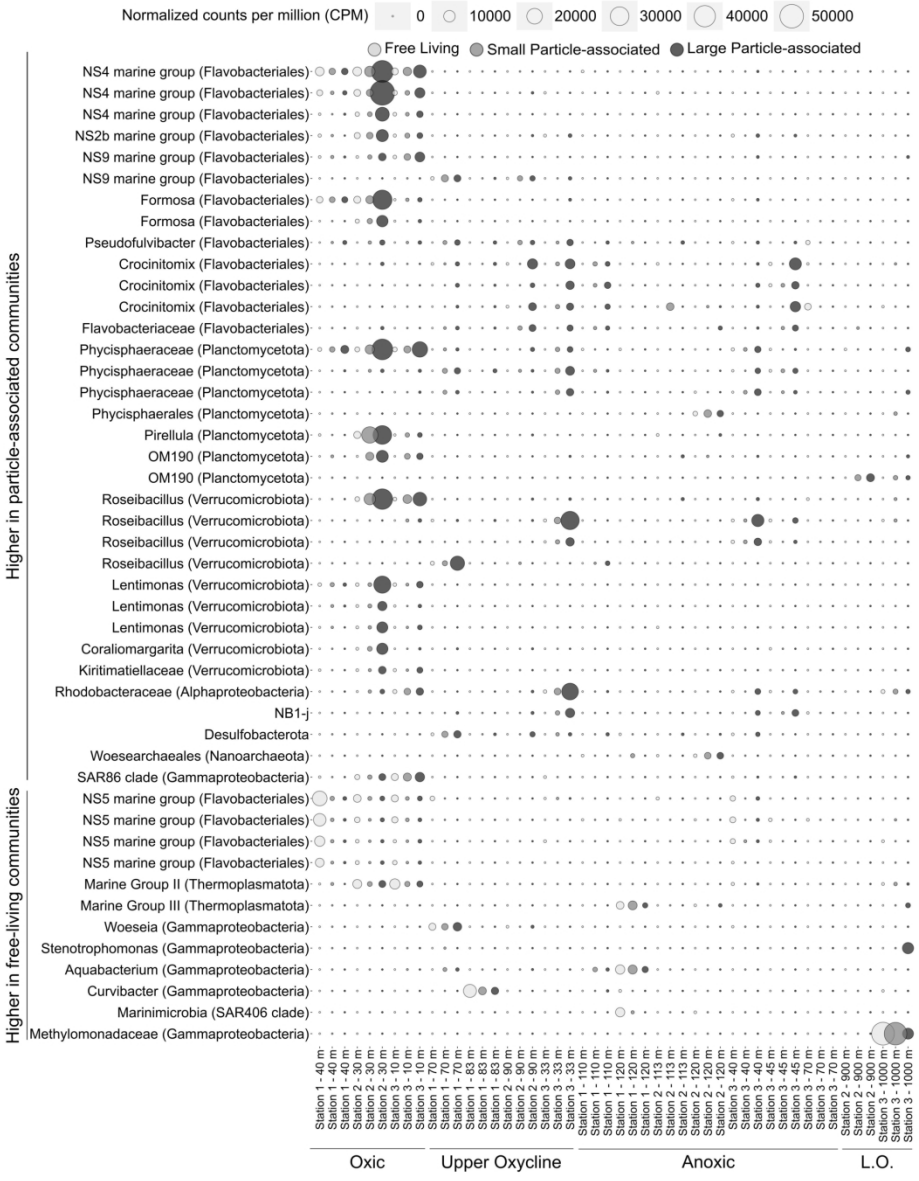

Figure 5. The normalized (by trimmed mean of M values) abundance (counts per million calculated by EdgeR) of 46 most abundant amplicon sequence variants (rows) that were differentially abundant between surface, upper oxycline and anoxic depths, and lower oxycline (L.O.) depths in all samples (columns) collected from the ETNP OMZ in the free-living (FL), small particle-associated (SPA), and large particle-associated (LPA) size fractions.

174x223mm (300 x 300 DPI)

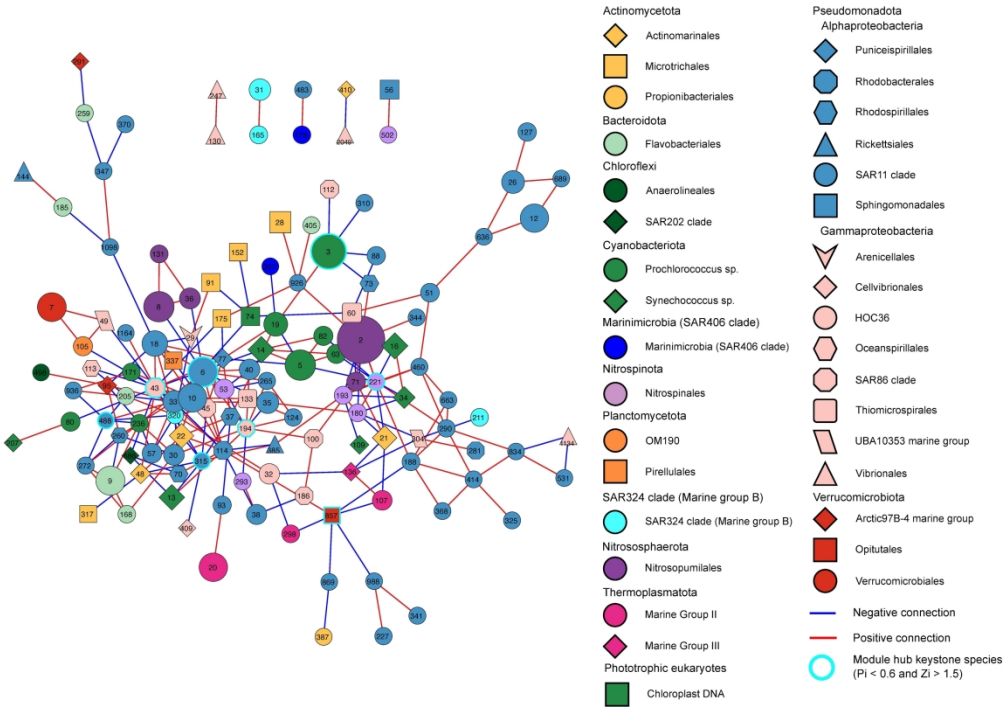

Figure 6: Co-occurrence network interactions of free-living microbes in the ETNP OMZ. Each node is labeled with an ASV code. The area of each node, which represents an amplicon sequence variant (ASV), is proportional to the non-transformed relative abundance of that ASV (the smallest node (ASV2049) has a relative abundance of 0.0000309%, and the largest node (ASV2) has a relative abundance of 0.0233%). Positive interactions are shown in red and negative ones in blue. A teal outline around the symbol highlights ASVs identified as putative keystone species (Zi > 1.5 and Pi < 0.6).

256x179mm (300 x 300 DPI)

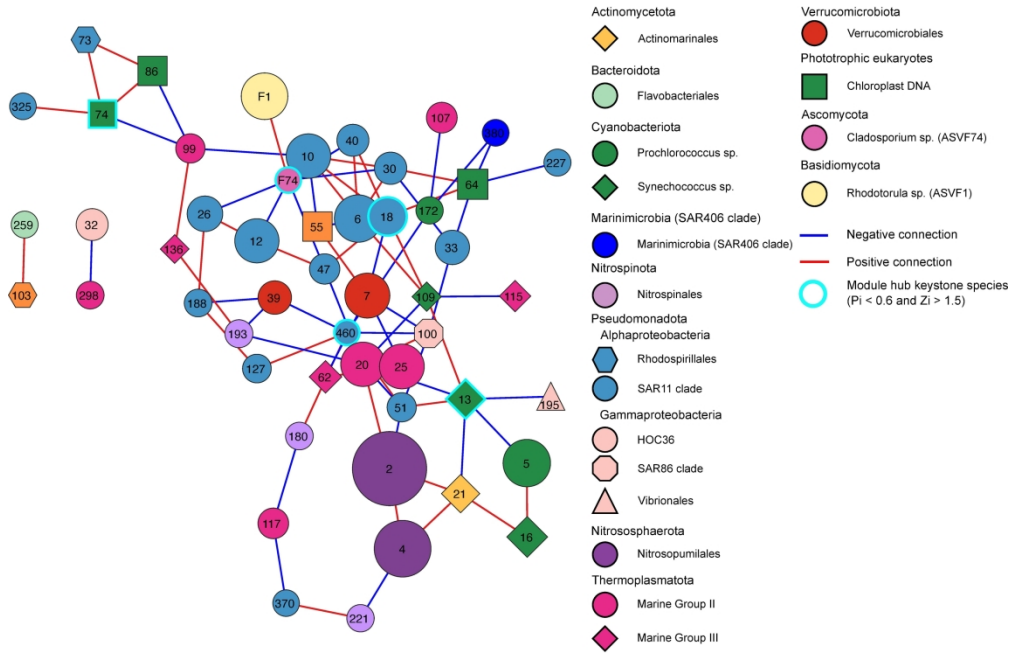

Figure 7: Co-occurrence network interactions of small particle-associated microbes in the ETNP OMZ. Each node is labeled with an ASV code. The area of each node, which represents a prokaryotic or fungal amplicon sequence variant (ASV), is proportional to the non-transformed relative abundance of that ASV (the smallest node (ASV370) has a relative abundance of 0.000348%, and the largest node (ASV2) has a relative abundance of 0.023%). Positive interactions are shown in red and negative ones in blue. A teal outline around the symbol highlights ASVs identified as putative network hubs (Zi > 1.5 and Pi < 0.6).

265x171mm (300 x 300 DPI)

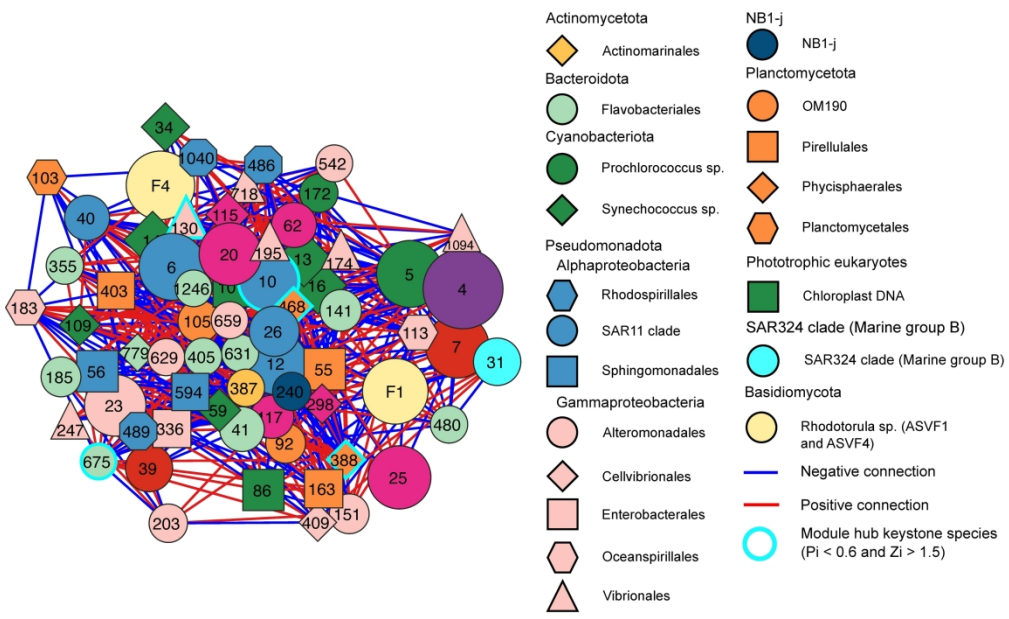

Figure 8: Co-occurrence network interactions of large particle-associated microbes in the ETNP OMZ. Each node is labeled with an ASV code. The area of each node, which represents a prokaryotic or fungal amplicon sequence variant (ASV), is proportional to the non-transformed relative abundance of that ASV (the smallest node (ASV1040) has a relative abundance of 0.0000956%, and the largest node (ASV4) has a relative abundance of 0.0148%). Positive interactions are shown in red and negative ones in blue. A teal outline around the symbol highlights ASVs identified as putative network hubs (Zi > 1.5 and Pi < 0.6).

298x178mm (300 x 300 DPI)