Microbial diversity and abundance vary along salinity, oxygen and particle size gradients in the Chesapeake Bay

Authors

Jacob A. Cram[1]*, Ashley Hollins[1], Alexandra J. McCarty[1, 2], Grace Martinez[3], Minming Cui[4], Maya L. Gomes[4], Clara A. Fuchsman[1]

[1]University of Maryland Center for Environmental Science - Horn Point Laboratory

Cambridge, Maryland, United States

[2] Virginia Institute of Marine Science - Marine Advisory Program; Gloucester, Virginia, United States

[3]Maryland Sea Grant College, College Park, MD 20740

[4]Johns Hopkins University - Earth and Planetary Sciences; Baltimore, Maryland, United States

*Correspondence: jcram@umces.edu

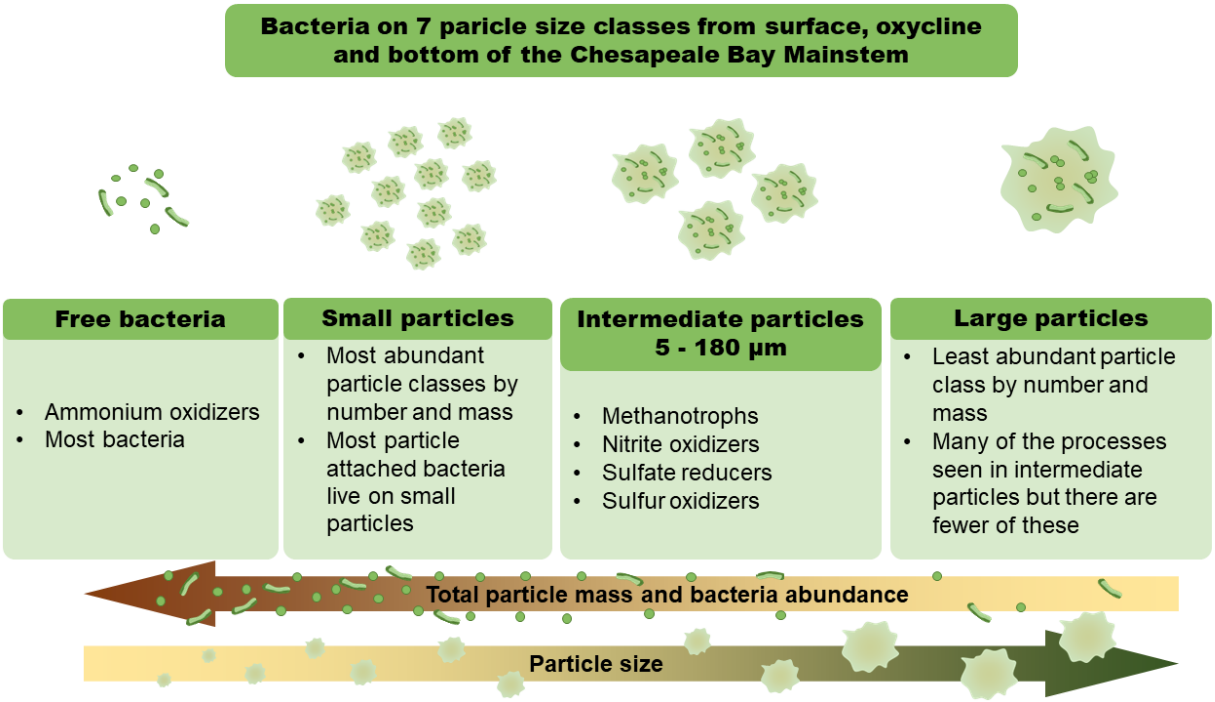
Abstract

Marine snow and other particles are abundant in estuaries, where they drive biogeochemical transformations and elemental transport. Particles range in size, thereby providing a corresponding gradient of habitats for marine microorganisms. We used standard normalized amplicon sequencing, verified with microscopy, to characterize taxon-specific microbial abundances, (cells per liter of water and per mg of particles), across six particle size classes, ranging from 0.2 to 500 μm, along the main stem of the Chesapeake Bay estuary. Microbial communities varied with salinity, oxygen concentrations and particle size. Many taxonomic groups were most densely packed on large particles (in cells/mg particles), yet were primarily associated with the smallest particle size class, because small particles made up a substantially larger portion of total particle mass. However, organisms potentially involved in methanotrophy, nitrite oxidation, and sulfate reduction were found primarily on intermediately sized (5 - 180 μm) particles, where species richness was also highest. All abundant ostensibly free-living organisms, including SAR11 and *Synecococcus*, appeared on particles, albeit at lower abundance than in the free-living fraction, suggesting that aggregation processes may incorporate them into particles. Our approach opens a door to a more quantitative understanding of the microscale and macroscale biogeography of marine microorganisms.

Abbreviated Summary

We examined bacteria that live on microscopic particles of six different sizes, and not on particles at the surface and bottom of six locations (some locations lacked oxygen) in the Chesapeake Bay. Most of the bacteria that live on particles live on the smallest particles, because those particles are the most abundant. However, some bacterial groups, especially ones that change the water's sulfur and nitrogen chemistry are often most abundant on particles of intermediate size.

Graphical Abstract



Introduction

Marine environments exhibit microscale heterogeneous environmental conditions and bacterial communities vary on the scale of microns to millimeters (Long and Azam 2001; Simon et al. 2002; Azam and Malfatti 2007; Stocker 2012). A main contributor to the ocean's microscale heterogeneity are particles that vary in size, shape, density, and chemical composition (Alldredge and Silver 1988). These particles are habitats for microorganisms whose metabolic and behavioral niches differ from those of free-living organisms (Simon et al. 2002; Leu et al. 2022). Particle attached bacteria experience environments with more abundant energy sources, harbor more diverse genes for breakdown of peptides and carbohydrates, and are more densely packed, than are free-living bacteria (Alldredge et al. 1986; Simon et al. 1990; Leu et al. 2022).

Not only are free-living microbial communities different than particle-attached ones (DeLong et al. 1993; Bidle and Fletcher 1995), there is also substantial variability between the microbial community on different particles (Mestre et al. 2017; Farnelid et al. 2018). The differences between microbial communities on particles of different size (such as on 20 – 200 µm vs 0.8 – 3 µm particles) are of similar magnitude to those on particles of the same size found at different layers of the deep ocean water column (such as the surface and the mesopelagic) or in different ocean basins (such as the Pacific and the Atlantic Oceans) (Mestre et al. 2018). These microbial community structures reflect chemical differences; particles of different sizes are believed to harbor different chemistry with larger particles harboring diffusion gradients of oxygen and other oxidants such as nitrate, which allow anoxic microzones to form in low-oxygen

water and sulfidic microzones to form in anoxic water (Ploug et al. 1997; Stief et al. 2016; Bianchi et al. 2018; Fuchsman et al. 2019b; Saunders et al. 2019; Raven et al. 2021).

The Chesapeake Bay is the largest estuary in the United States and provides a well-studied model system characterized by high production and active biogeochemical processes (Turk et al. 2021). The Bay is characterized by high particle abundance, which transport nutrients and carbon through the system (Sanford et al. 2001; Malpezzi et al. 2013; Palinkas et al. 2019). Particles transport organic carbon to the middle of the Bay, where it fuels microbial respiration, depleting the mid-Bay of oxygen (Wang and Hood 2020) and creating a seasonally (summer) oxygen-starved environment (Testa et al. 2018). Bacteria in the anoxic Bay are known to produce greenhouse gases including methane (Gelesh et al. 2016) and nitrous oxide (Ji et al. 2018; Laperriere et al. 2019), as well as hydrogen sulfide (Luther et al. 1988), which is toxic to marine life (Kang 1997; Boyd 2014). Sulfur oxidizing bacteria, responsible for the removal of hydrogen sulfide and other reduced sulfur species have been identified in the hypoxic Bay (Crump et al. 2007; Findlay et al. 2015), potentially using nitrogen species as terminal electron acceptors (Arora-Williams et al. 2022). Methane appears to be produced in the sediments, but is oxidized in the water column (Reeburgh 1969; Hagen and Vogt 1999; Gelesh et al. 2016). However, it is unknown whether and on what sizes of particles methane and sulfur cycling bacteria associate. Bay microbial communities vary across space and season, following the changes in oxygen/sulfide concentrations (Kan et al. 2006, 2007; Wang et al. 2020; Arora-Williams et al. 2022). However, no investigation of the spatial variability of particle associated bacterial communities using modern techniques has been done in the Bay.

Highly size resolved measurements from global sampling efforts have shown that particle size variability promotes microbial diversity globally (Mestre et al. 2018). Previous analyses of bacterial communities along the particle size spectrum have been semi-quantitative, providing relative abundance estimates (fraction of the total community), rather than estimates of quantitative abundance (cells per liter of water or milligram of particle). Therefore, we implemented a novel size fractionation approach, with seven size fractions, to allow collection of DNA, microscopic samples, and concentrations of the particles themselves in each size fraction. We furthemore utilized DNA standards to quantitatively describe microbial distributions along the particle size spectrum. Here, we describe particle size resolved measurements of microbial communities and how they vary across space in the Chesapeake Bay.

Methods

Samples were collected at six stations along the length of the mainstem of the Chesapeake Bay (Figure 1A) at the surface and the bottom of the water column between 2019-July-22 and July-24. Samples were collected at Chesapeake Bay program stations CB3.1 and CB3.2 at depths of 3 m (Surface) and 7 m (Bottom); CB3.3 at 3 m (Surface) and 7 m (Oxycline); station CB4.3C at depths of 3 m (Surface) 6 m (Oxycline) and 19 m (Bottom); Station CB 5.1 at 7 m (Surface) and 32.5 m (Bottom); and Station 5.5 at 3 m (Surface) and 13 m (Bottom). As stations were located in 13.3, 12.2, 24.1, 27.1, 34.3 and 17.7 m of water respectively, all samples labeled “bottom” were taken near the bottom of the water column. Samples were collected on July 22 from stations CB5.1 and CB5.5, July 23 at stations CB4.3C and CB3.3C and July 24 at stations CB3.2 and CB3.1. Hydrological conditions were assessed during the time of sampling by querying Chesapeake Bay Program stations (see supplementary results).

Approximately 15 L of water was collected per station using Niskin bottles on a shipboard CTD rosette. Water was removed from the Niskin bottles by opening the lower stopper in order to collect even those particles that settled below the sampling valve. We collected POM from six size classes: 500 µm and larger, 180 – 53 µm, 53 – 20 µm; 20 – 5 µm and 5 – 1.2 µm. We collected DNA from all of the above size fractions and a 1.2 – 0.2 µm size fraction. Sample processing happened in two phases. In the first phase, particles were size fractioned using nylon mesh and re-suspended into a particle slurry of particulate matter made of particles from 500, 180, 53, 20, 5 µm size classes. Additionally, during this stage, water containing only particles smaller than 5 µm and free-living bacteria was saved. In the second phase, particle slurry from each size larger than 5 µm was collected on filters for analysis of particulate matter content (GF/C) and molecular analysis (Supore), and for microscopy (formalin preserved water). Additionally, during the second stage, water that had passed through the 5 µm filter was split, with half passed through 1.2 and 0.2 µm Supore filters in series for collection of DNA and a 1.2 µm GF/C filter for collection of POM. Additionally, a portion of this water was also preserved for microscopy analysis. A full description of these two phases can be found in the supplement (Supplemental Methods; Particle Processing).

Particulate Organic Matter Mass

To measure particulate organic matter mass, GF/C filters were post-weighed and the pre- and post-weights were compared, as described and reported in Dougherty et al. (2021). Total organic matter mass per sample was normalized to the volume of water filtered through the nylon filter and the fraction of rinse water that passed through the GF/C filter. Total particulate matter mass was calculated following Eqn. 1.

$$Particle\,mass(mg/L) = \frac{FilterPostWeight(mg) - FilterPreWeight(mg)}{Volume\,Filtered(L)} * \frac{Total\,Volume\,of\,Rinse\,Water(L)}{Volume\,Rinse\,Water\,Used\,for\,POM\,Measurement(L)} \quad \text{Eqn. 1}$$

Isotopic Analysis

After mass was measured, GF/C filters were wafted with hydrochloric acid vapor for 24 hours to remove carbonates, dried, packed in both silver and tin capsules, and sent to University of California Davis Stable Isotope Facility. Blank combusted GF/C filters were also included in the analysis. Samples were processed for mass spectrometry following their Difficult Combustion of Solid Samples Protocol (Supplemental Methods; Isotopic Analysis). Total particulate nitrogen and particulate carbon concentrations were converted into concentrations by modifying Eqn. 1, substituting the filter mass difference with observed carbon and nitrogen concentrations.

Hydrogen Sulfide Measurements

Samples for hydrogen sulfide concentration analysis were treated with zinc acetate to stabilize sulfide and stop biological activity. Hydrogen sulfide concentrations were measured at stations CB3.3C and CB4.3C using a colorimetric assay, separately at Horn Point Laboratory (Station CB4.3C) and at Johns Hopkins University (all depths at station CB3.3C and 21.9m at CB4.3C) (Cline 1969; Parsons 1984).

Measuring Microbial Diversity and Abundance

Microbial abundance on selected samples was measured from the formalin preserved particles by removing the bacteria from the particles by adding detergent and sonicating, and then enumerating bacteria using DAPI based autofluorescence microscopy (Supplemental Methods: Microscopy measurements of bacterial abundance). DNA was extracted from the Supore filters using an in-house phenol chloroform process (modifying Fuhrman et al. 1988; Cram et al. 2016) (Supplemental Methods; DNA Extraction). Prior to amplification, 10^5 copies of a synthetic 16S rRNA sequence that has an identity distinct from any known organism (GenBankAccession Number LC120931; Tourlousse et al. 2017) was added per ng of environmental DNA. DNA was amplified with slight modifications to Needham et al.'s published protocol (2018) (Supplemental Methods; Amplicon Libraries). Amplicon sequence variants were called using the DADA2 algorithm (Callahan et al. 2016), following a modified version of Lee et al.'s (2019) protocol (Supplemental Methods; Amplicon Bioinformatics).

Taxon Specific Abundance Estimation

To estimate the environmental abundance of each ASV, each ASV sequence read count was normalized to spike in read counts, the total amount of DNA extracted from each sample, volume of water filtered, and rinse water volume following Eqn. 2.

$$Taxon\ Abundance(16S+18S\ copies/L) = \frac{ASV\ Reads}{Spike\ Reads} * \frac{10^5 Spike\ Copies}{1ngDNA} * \frac{DNA\ Extracted(ng)}{Volume\ Filtered(L)} * \frac{Total\ Volume\ of\ Rinse\ Water(L)}{Volume\ Rinse\ Water\ Used\ for\ DNA\ Extraction(L)} \quad \text{Eqn 2.}$$

Taxon abundance was further normalized to the width of the particle size fraction bins (Eqn 3.)

$$NormalizedTaxonAbundance(copies/L/\mu m) = \frac{Taxon\ Abundance(copies/L)}{SizeClassUpperBound(\mu m) - SizeClassLowerBound(\mu m)} \quad \text{Eqn 3.}$$

To estimate microbial cells per mg of particle mass, microbial abundance was normalized to particle mass (Eqn. 4).

$$Taxon\ Abundance(copies/mg) = \frac{Abundance(copies/L)}{Particle\ mass(mg/L)} \quad \text{Eqn. 4.}$$

Analytical approach

Alpha diversity

We used the `breakaway` package (Willis et al. 2018) to estimate species richness in each of our samples and used the `betta` function therein to explore how richness varied with latitude, depth and particle size. We used a polynomial model, in which we included a squared term for latitude and particle size to identify whether richness was highest or lowest at intermediate salinities and particle sizes.

We calculated the Shannon index (H) of diversity using the `vegan` package, first rarefying samples to 806 reads, which was the lowest number of reads in any of the samples that were included in the analysis. We estimated Pielou's evenness index (J), by dividing the Shannon index (H) by breakaway's estimate of evenness (ignoring breakaway's confidence intervals). We

applied ordinary least squares linear models, using R's base `stats` package to estimate the relationship between the Shannon diversity and Pielou evenness scores and particle size, salinity and oxygen, again including polynomial terms for salinity and depth. For consistency we also used the linear model to estimate how the richness estimates (from `breakaway`) varied with size, latitude and depth. Thus between the `betta` function (Willis et al. 2018) and the linear model, we had two different models of how richness varied with size, salinity and oxygen.

Beta Diversity

Overall community patterns and their relationship to sample latitude, depth and size were summarized by using the `rda()` function in the `vegan` package (Oksanen et al. 2013) on log-transformed volume and bin size normalized microbial abundance values (cells/L/μm). Significance testing was performed using a permutation test for redundancy analysis as implemented in vegan's `anova.cca()` function. To reduce computational complexity and challenges from zero inflation, we only included ASVs in the analysis that appeared in at least 20% of samples.

Community structure

Abundance of microbial taxa, grouped to different taxonomic levels were visualized and representative examples shown herein. We estimated Phylum level abundance patterns by summing the abundance of all ASVs within each phylum and then reporting only those phyla that comprised at least 10^6 copies/mg particles total, in any of the samples from the 1.2 μm or larger samples. To explore patterns within one phylum, we visualized all ASVs within the Planctomycetes phylum considering only ASVs that comprised at least 10^6 copies/mg of particles in the particle containing samples. Planctomycetes were chosen as the representative phyla for three reasons: (1) they are known to be a major clade of bacteria that are predominately particle associated (DeLong et al. 1993; Fuchsman et al. 2012), (2) they are abundant and widespread in the Bay (Kan et al. 2006), and (3) they are important players in the marine carbon and nitrogen cycles (Shu 2011). We also visualized the abundance of those bacteria that were most abundant on particles 20 μm or larger. In this case we only showed ASVs that had an abundance of at least 10^3 copies/mg particles on any one sample.

Estimating biogeochemical function

To identify bacteria involved in sulfur cycling and methanotrophy, we used tools from `PICRUSt2` (Douglas et al. 2020) to identify which ASVs had most closely sequenced relatives that harbored genes for methanotrophy (particulate monooxygenase; EC:1.14.18.3) and for sulfur cycling (dissimilatory sulfite reductase; EC:1.8.99.5) from the Kegg EC enzyme database (Kanehisa 2017). From within the `PICRUSt2` package, we aligned the sequences with a reference tree using `HMMER`, found the most likely placements of each ASV on that reference tree with `EPA-NG` (Barbera et al. 2019) and output a tree file with `GAPPA` (Czech et al. 2020), using `PICRUSt2`'s `place_seqs.py` command. We then used the `hsp.py` function that implements the `castor` algorithm (Louca and Doebeli 2018) to predict gene families associated with each ASV. While the pipeline for `PICRUSt2` conventionally extends to predicting metagenomic potential of the overall community, we stopped after gene prediction, and instead identified which ASVs were associated with our two genes of interest.

To identify bacteria putatively involved in nitrogen cycling, we identified ASVs with in which any taxonomic identifier, from class to genus level, began with the string “Nitro-”, and confirmed that these were known ammonium and nitrite oxidizing organisms. Similarly to identify bacteria involved in methaotrophy, we searched for archaea with taxonomic identifiers beginning with the string “Methano-” (Garcia et al. 2000) and for archaea from the Verstraetearchaeota phylum (Vanwonterghem et al. 2016). We visualized the abundance of these particularly biogeochemically interesting ASVs, including only ASVs that have an abundance of at least 10^5 copies/L in at least one size fraction.

Results

Site description

As reported previously (Dougherty et al. 2021), stations follow a salinity gradient with northernmost stations less saline than more southerly stations. All sites are characterized by a pycnocline between 5 and 10 m in depth, with a sharp oxycline at the central Bay stations CB3.3 and CB4.3C (Figure 1A-C). There was an observed gradient of hydrogen sulfide at stations CB3.3C and CB4.3C, which was evident from a sulfide smell in the water. A detailed profile of hydrogen sulfide was measured at CB3.3C and CB4.3C with a notable increase below 15 m at both stations (Figures 1C, S3). Hydrography corresponding to our times of sampling are described in supplement.

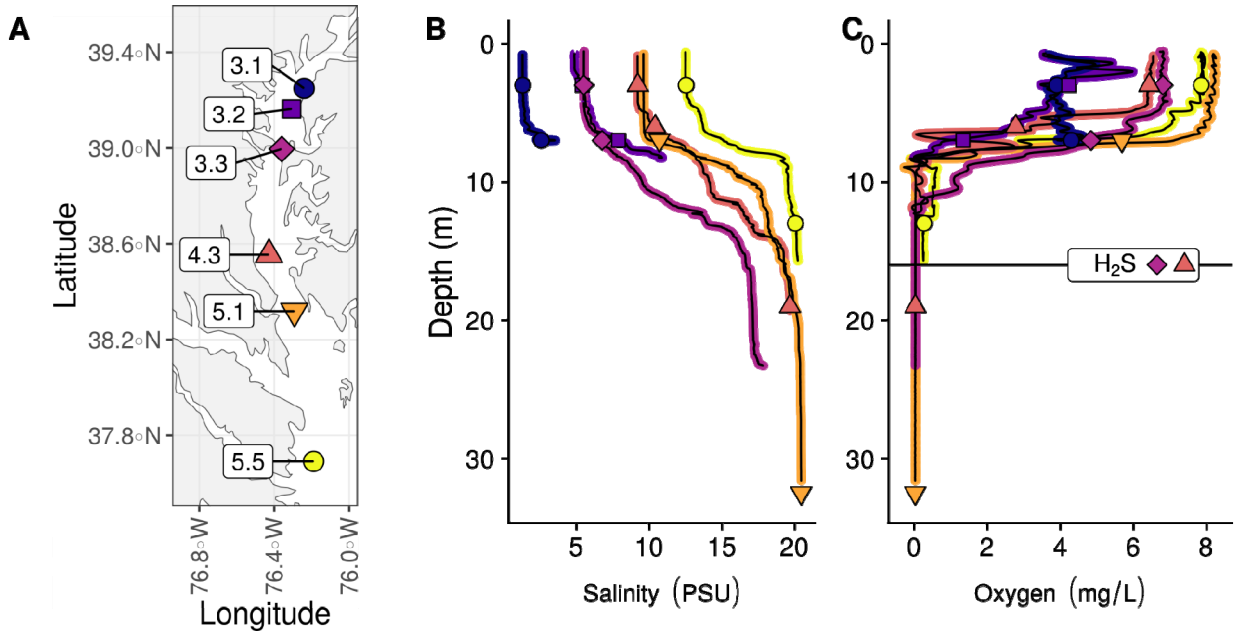


Figure 1. Description of the physics and chemistry of the sample sites **A.** Map of all sample sites – color and shape coding of sites corresponds to point color and shape in panels B and C. Stations correspond to the Chesapeake Bay Program sampling grid, all station names begin with the prefix “CB,” and stations 3.3 and 4.3 correspond to central Bay stations and end with the suffix “C”. For instance, “4.3” corresponds to station “CB4.3C.” **B.** Vertical profiles of Salinity and **C.** Oxygen. The horizontal line in *C,* labeled with H_2S, indicates the depth at which hydrogen sulfide exceeds 16 µM at station CB3.3C and where the water is also sulfidic at

CB4.3C (Figure S3). In *B* and *C* shape position of each point along the depth axis shows the depth of microbial sampling.

Particle composition

As reported previously, particle mass follows an inverse power law relationship with particle size (Dougherty et al. 2021) (Figure 2A). Particle carbon and nitrogen mass both also follow power laws with respect to size (Figure S4A-C). Generally, carbon comprises ~3 – 30% and nitrogen ~0.3 – 3% of total particle mass, regardless of particle size (Figure S5AB). Carbon to nitrogen ratios vary between environments, with highest relative nitrogen content at the northernmost stations CB3.1 (C:N, integrated over all size classes Surface = 9.6, Bottom = 9.2) and CB 3.2 (C:N Surface = 8.6, Bottom = 7.8), and highest relative carbon content at station CB4.3C (C:N Surface 24.3, Oxycline = 16.0, Bottom = 29.0) but are consistently well above the Redfield ratio (C:N = 6.6; Figure S5C), indicating degraded material. $\delta^{13}C$ appears to vary latitudinally at the surface, running between -23 and -30‰ with least negative values at the southernmost stations, and most negative values at the terrestrial stations (Figure S6A), as would be expected due to the isotopic composition of terrestrial (more depleted) and marine (less depleted) organic matter (Arthur et al. 1985). In bottom waters, the least negative values appear to be found at the intermediate stations, though these patterns vary between size classes. $\delta^{15}N$ values appear to usually be around 10‰ with some specific samples having higher values (Figure S6B).

Microbial Total Abundance

Amplicon sequences suggested acceptable sequence quality data, as evidenced by sequencing of mock communities (Supplemental Results). Microbial 16S and 18S gene abundances, generated by this approach, were generally in the range of 10^7 - 10^8 copies per mg of particle mass, regardless of particle size or location (Figure 2D). The exception was samples taken at the northernmost, least saline station CB3.1. There, microbial abundance on intermediate sized particles (5 – 53 μm) was between 10^5 – 10^6 copies per mg. We expect these abundances are likely slight (same order of magnitude) over-estimates of total microbial abundances because single cells often harbor multiple 16S or 18S gene copies (Větrovský and Baldrian 2013). Consistent with this, amplicon-based estimates of microbial abundance were generally within the same order of magnitude as, though slightly higher than, microscopy-based estimates (Figure S8), as expected from organisms that harbor multiple 16S and 18S gene copies per cell.

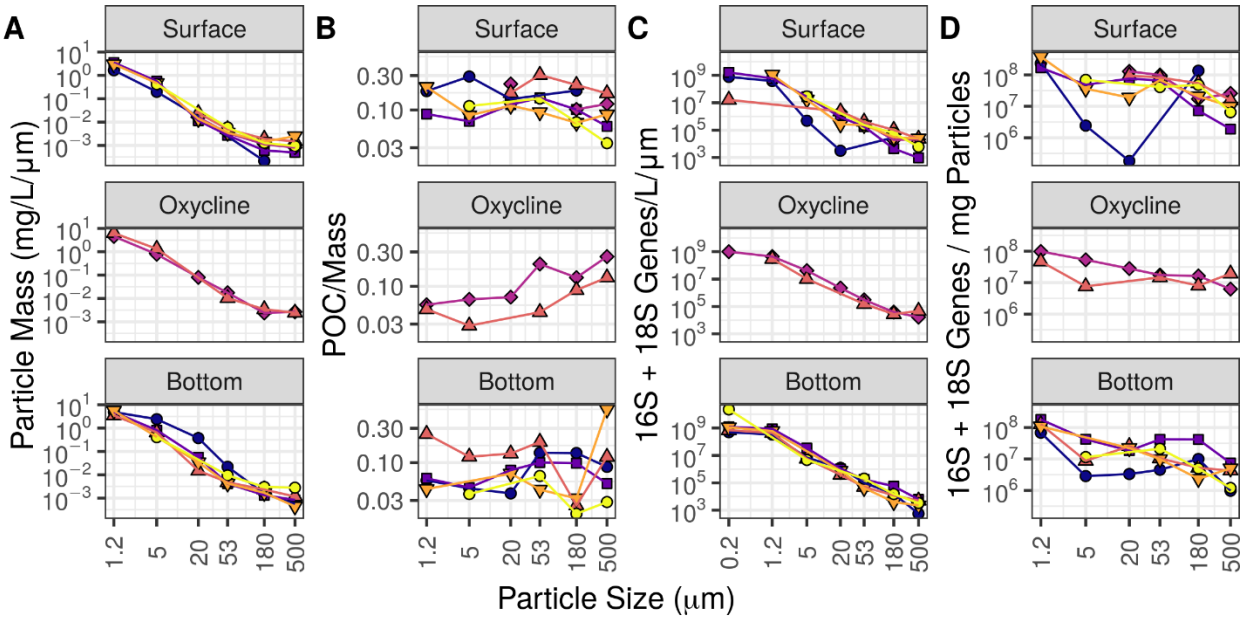


Figure 2. Free living and particle associated microbial total abundance at each station and depth. **A.** Total particle mass, normalized to particle size bin width. **B.** The ratio of particulate carbon mass to total particle mass. **C.** Microbial abundance per liter of water, normalized to particle size bin width. **D**. Microbial abundance normalized to particle mass. In all panels, both X and Y axes are on a log scale. Line and symbol colors indicate stations as seen in Figure 1.

Alpha diversity

Across our entire dataset, we observed 82476 unique ASVs, excluding ASVs that mapped to our spike-in samples. The breakaway package, which estimates the abundance of unobserved singleton species, estimated that sample richness ranged from 23 (lower bound 22.1, upper bound 39.3) to 10.2 * 10^3 (bounds: 4.90 * 10^4 – 2.77 * 10^5) ASVs per sample (Figure S9A). The `betta` function, which accounts for the abundance of these unobserved species on richness patterns (Willis et al. 2018) showed non-linear relationships with particle size and salinity ($p < 0.001$; Figure S10A; Table S4), with highest richness among intermediate sized particles (5–20 µm and 20–53 µm size bins; Figure S10A). Richness appeared to be unrelated to dissolved oxygen concentration (Table S4; $p = 0.779$). Applying a simpler linear model, rather than the `betta` function showed similar but weaker patterns (Table S4). Thus, for our data, the `betta` algorithm is consistent with, though more sensitive than, the linear model.

Linear models further suggested that the Shannon diversity index (Figures S9B, S10B) was, like richness, highest among intermediate particle sizes (5 – 20 µm, 20 – 53 µm, and 53 – 180 µm size bins; Table S4). Evenness (Figures S9C, S10C) was non-statistically significantly lowest among intermediate sized particles ($p = 0.053$) and did not vary with any other parameters (Table S2).

Beta Diversity

Redundancy analysis suggested that overall microbial community structure (normalized to volume filtered and bin-size, and expressed as copies/L/µm) was statistically related to particle

size, water salinity and water oxygen concentrations (Figure 3; Figure S11; Table S3; ANOVA $p < 0.01$ for all terms).

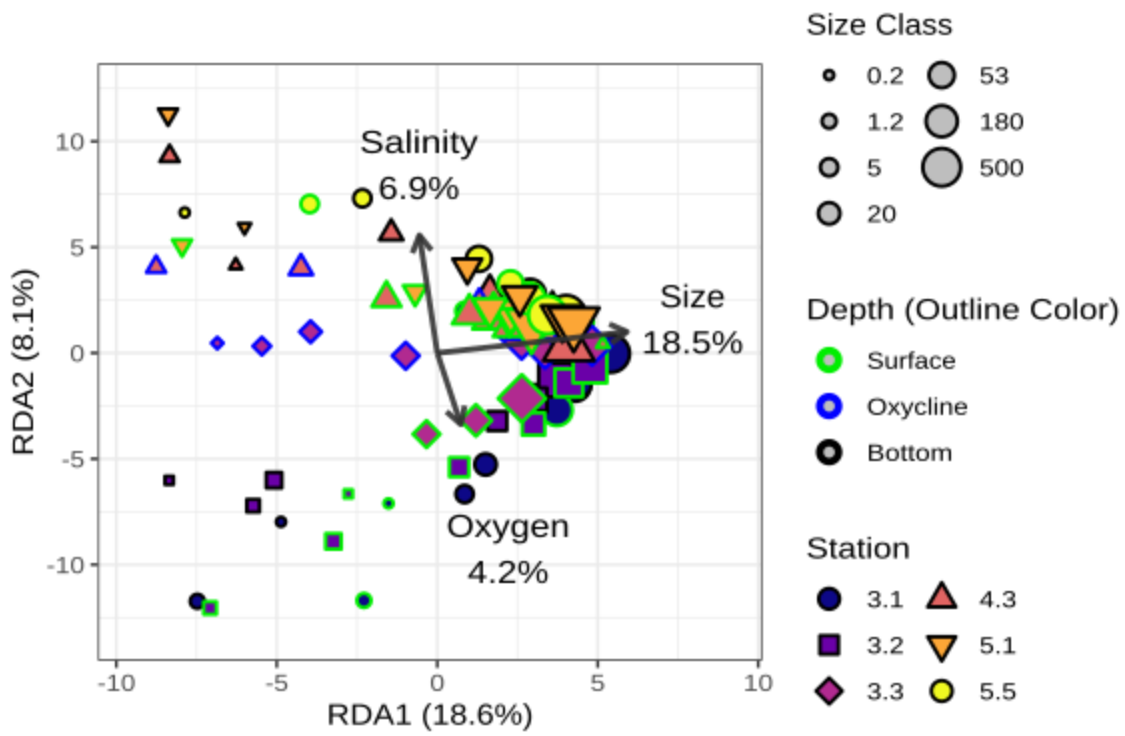


Figure 3. Redundancy analysis of the relationship between community structure and salinity, oxygen concentration and particle size. Both size and oxygen data have been log transformed. Axes correspond to the first two redundancy analysis axes and show 18.6% and 8.1% of the variance. Points indicate samples and their outline colors, fill colors and shapes (corresponding to stations indicated in Figure 1A) and sizes indicate which station, depth and size class they were collected from (legend). Arrows indicate the three terms in the RDA model: particle size, salinity, and dissolved oxygen concentrations. Percentages by each arrow show the marginal percentage of variance explained by that parameter.

Similar patterns were evident when samples from the 0.2 – 1.2 µm fraction were removed, and when samples from the 0.2 – 1.2 and 1.2 – 5 µm fractions were both removed (Figure S12; Table S5), indicating that community structure differs across the particle size gradient and is not driven only by differences between free-living and attached bacteria. Similar patterns were also observed, when latitude, depth (whether the sample was collected from near the surface) and size (hereafter Positional Model) were compared to community structure (Figure S13A,B), rather than salinity, oxygen and size (hereafter Environmental Model). In both cases, the predictor variables explained a similar fraction of total variance (Environmental Model 26.3%; Positional Model, 25.4%). We found that adding sample depth to the Environmental Model did not result in

an increase in model performance (ANOVA; $p = 0.31$). For both models, adding quadratic terms lead to an increase in model performance, with lower AIC values (Environmental Model, DF = 4, AIC = 681; Quadratic Environmental Model, DF = 6, AIC = 677; Positional Model, DF = 4, AIC = 682; Quadratic Positional Model, DF = 6, AIC = 677) and higher percent variance explained (Quadratic Environmental, 34.3%; Quadratic Positional, 34.1%). Adding information about carbon to mass ratio, carbon to nitrogen ratio, $\delta^{13}C$ or $\delta^{15}N$ values of the particles to the Environmental Model did not appear to lead to a statistically detectable increase in model performance. (ANOVA; p of all new terms > 0.05). Indeed, a stepwise regression in which non-significant terms were dropped from the model in order of lowest significance found that each of these new terms was eliminated from the model.

Diversity Patterns

Phylum Level

It was clear that some microbial taxa were associated in particular with some stations, depths and particle size fractions. Normalized to particle mass, bacteria from the phyla Bacteroidetes, Cyanobacteria, Planctomycetes, Verrucomicrobia and Alpha-, Delta- and Gamma- Proteobacteria were all abundant on particles, each with higher abundance at the surface of all stations, the bottom of station CB3.2, and the oxycline of station CB3.3C (Figure 4). Archaea in contrast were rarely abundant on particles. Euryarchaeota were abundant on small particles at station CB3.2 Bottom and CB4.3C Oxycline. Diverse eukaryotes, especially meso-zooplankton and Ochrophyta (diatoms and other brown algae) were detected and abundant at least at some stations. All were associated primarily with intermediate and large size classes, likely reflecting the larger size of these organisms.

Normalized to water volume, rather than particle mass, all particle associated phyla were primarily associated with the smallest particles (Figure S14). This is because even if the microorganisms are abundant relative to particle mass, most particle mass is associated with small particles (Figure 2AB).

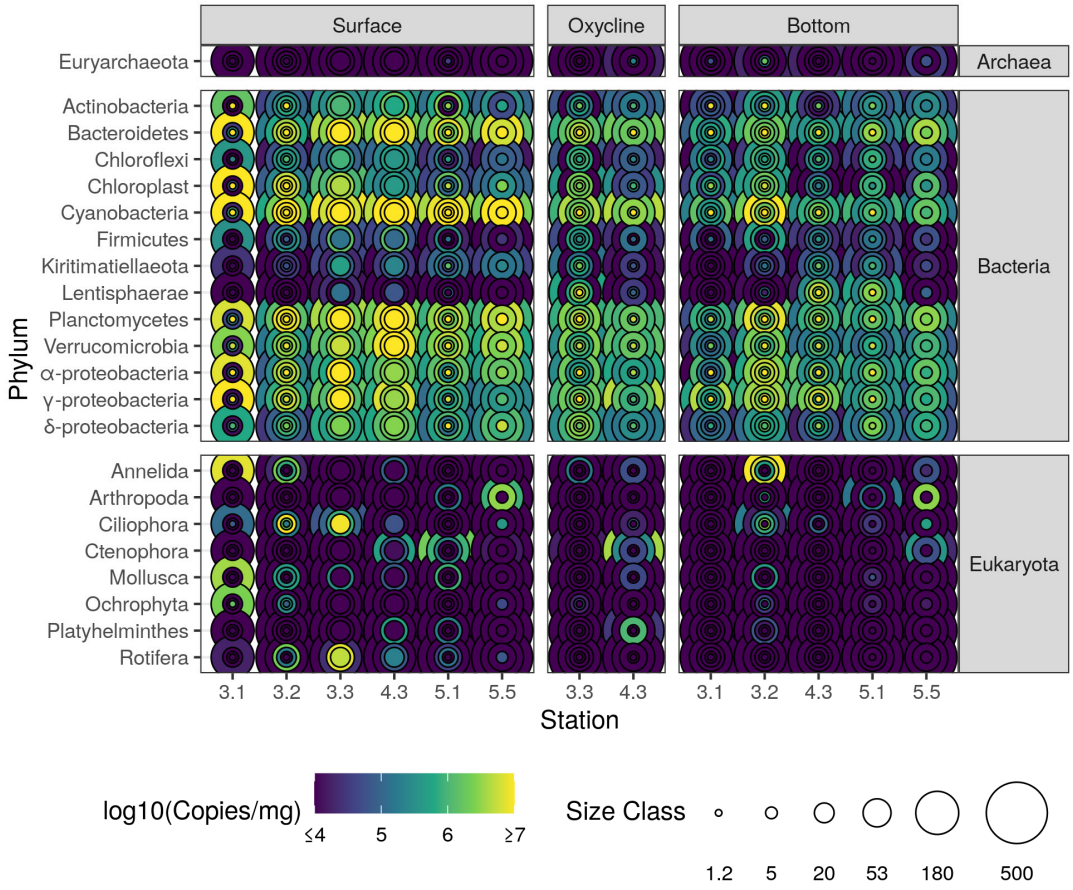


Figure 4. Phylum level taxonomic groups, measured as 16S or 18S rRNA gene sequences per mg of particulate mass, in all three domains of life. Phyla are in rows, with kingdoms shown in the panels at right. Stations are in columns with panels at top showing sample depth (Surface, Oxycline, Bottom); latitude for each station is shown in Figure 1A. Concentric circles indicate each size class of particles. Color corresponds to log transformed abundance of each microbial group. Some size classes at some stations are not shown, because either particle mass or amplicon measurements were not successful. Only phyla whose abundance exceeds 10^6 cells/mg particles, in at least one sample, are shown. The Proteobacteria phylum was subdivided into class, and Chloroplasts, while technically Cyanobacteria under the SILVA taxonomic scheme, are treated as their own phylum.

ASV Level

As with the phyla level patterns, some ASVs were abundant when normalized to particle mass, but were scarce when normalized to water volume. This pattern occurs because large particles are scarce relative to small ones (Dougherty et al. 2021) and so comprise less habitat. An example of this pattern is the ASV level groups of Planctomycetes, which fall into three families (Phycispareles, Pirellulales and Planctomycetales). ASVs showed associations especially with high or low latitude stations, and some with surface or bottom waters. Some were primarily associated with small particles, and others with all sizes of particles (Figure 5A).

However, normalized to water volume, all Planctomycetes ASVs were primarily associated with the free-living and 1.2 – 5 μm size fractions (Figure 5B).

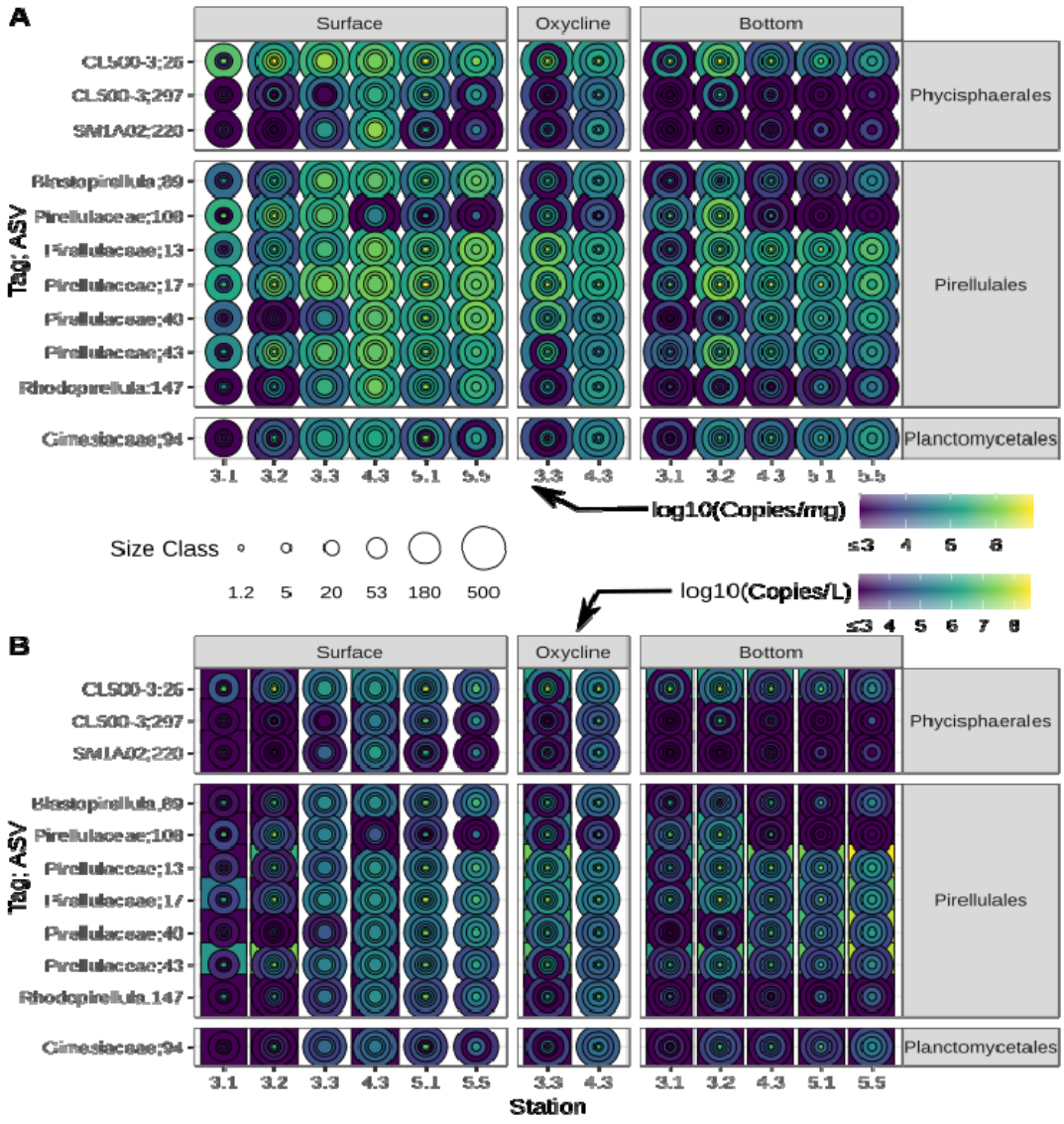


Figure 5. Abundance of different amplicon sequence variants from within the phylum Planctomycetes. ASVs are rows, with order level taxonomy on the panels at right. Depths and stations are indicated by column, as in Figure 4. **A.** Normalized to particle mass. **B.** Normalized to volume of water. Axes as in Figure 4. Only ASVs whose abundance exceeds 10^6 cells/mg of particles, in at least one sample, and that appear in at least 20% of all samples, are shown. ASV number is indicated after a semicolon.

A few ASVs showed an exception to the pattern in which most were associated with small particles. Six ASVs from the Alphaproteobacteria class and two each from the

Gammaroteobacteria class and Firmicutes phylum were primarily associated with size classes 20 µm or larger (Figure 6), were at least 10^6 cells/ml in one sample, and were observed in at least 20% of all samples. These species included known parasites such as an ASV from the Midichloraceae family, as well as others known to break down particulate matter such as one ASV from the Paracoccus genus, which BLAST against the NCBI database indicated was identical to two species: *Marcusii* and *P. Carotiniaciens*. Also evident in the larger size fractions were larger organisms. These included an ASV from the Lobata order (ctenophore) and a chloroplast for which NCBI Blast search reported 100 percent similarity to several eukaryotic algae, including both diatoms and foraminifera.

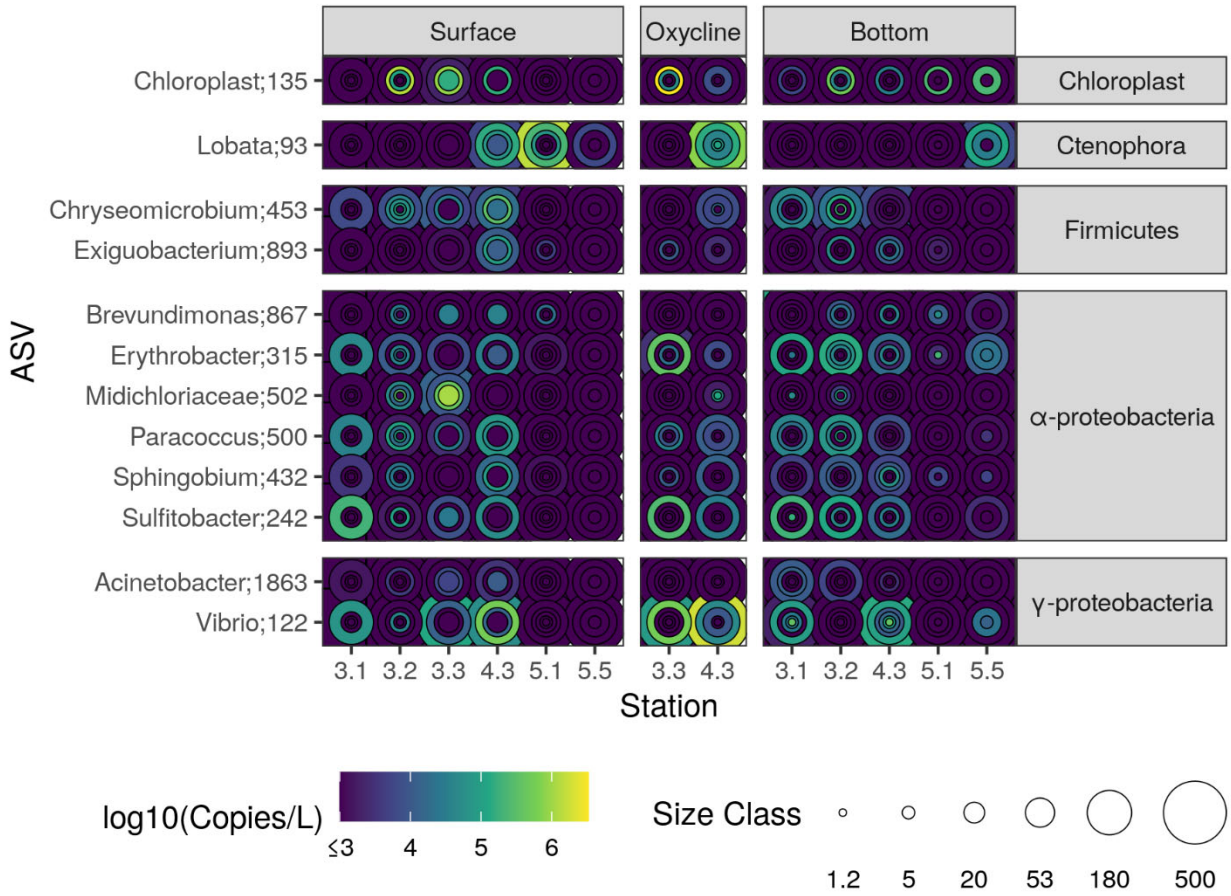


Figure 6. Abundance of different bacterial amplicon sequence variants that are most abundant, on average, on size fractions 20 µm or larger, normalized to volume of water. Axes as in Figure 4. Only ASVs whose abundance exceeds 10^4 cells/mg of particles, in at least one sample, and that appear in at least 20% of all samples, are shown. ASV number is indicated after a semicolon.

All bacterial ASVs that were abundant (≥ 1% of the total community) in the 0.2 – 1.2 µm, free-living size fraction were also found at lower, but still detectable abundance, in larger size fractions (Figure S15A-C). These included several bacterial ASVs from the SAR11 clade, which is ostensibly free living.

Bacterial influences on biochemical cycling

A range of bacteria potentially involved in methane, sulfur and nitrogen cycling, were identified by observing taxonomic identities, in the case of methanogenic and nitrogen cycling organisms; and using PICRUSt2 to identify bacteria with closest known relatives with that harbored the gene for dissimilatory sulfite reduction enzyme, in the case of sulfur cyclers; and the particulate monooxygenase enzyme in the case of methanotrophs (Figure 7). All of these biogeochemical processes were dominant in the bottom or oxycline samples, when they were observed.

Methane cycling

Methanotrophy: The PICRUSt2 based approach identified one species whose closest fully sequenced relative harbored the particulate monooxygenase enzyme EC:1.14.18.3, which is involved in methanotrophy with oxygen. This was an otherwise unidentified ASV from the Methylomonaceae family. This putative methanotroph was abundant particularly in the smallest particle size fraction (1.2 – 5 µm) in the bottom waters of all but two stations (Figure 7). At station 3.2 it was abundant on 20 – 53 µm and 53 – 180 µm particles, rather than the smallest size fraction. It was not abundant in any size fraction at the southernmost station CB5.5.

Methanogenisis: We identified one ASV that was potentially methanogenic, an unidentified ASV from the Methanofastidiosales order. This ASV was only found in one sample, the free-living (0.2 – 1.2 µm) size fraction at the bottom depth of station CB5.1 at an abundance of 1.6 x 10^5 16S gene copies/L seawater. No other groups were identified in our dataset with the prefix of Methano- in any of their taxonomic identifications from Class through Genus level (Garcia et al. 2000), nor were any bacteria from the Verstraetearchaeota phylum (Vanwonterghem et al. 2016) identified. This absence suggests that no other methanogens were present in our dataset.

Nitrogen Cycling

Ammonium oxidizing and nitrite oxidizing taxa were both evident at our site and were primarily found in bottom waters. Ammonium oxidizing archaea from the genus *Candidatus* Nitrosopumilus were found free-living in all stations except for the northernmost CB3.1 and most ASVs were also associated with the smallest particle size fraction at most of these sites. Ammonium oxidizing bacteria, one each from the genera *Nitrosoglobus* and *Nitrosomonas* genera were seen, with the *Nitrosomonas* found at the northernmost stations CB3.1 and CB3.2, again primarily free-living and on the smallest particle size fraction. *Nitrosoglobus* was found only in the free-living fraction at station CB5.5 and in no other samples. Nitrite oxidizing bacteria, all of the genus *Nitrospira*, showed a different pattern, in which they were primarily associated with intermediate sized particles (5 –20 µm and 20 – 53 µm) at stations CB3.1 and CB3.2, though one ASV was abundant on particles at station CB3.2.

Sulfur Cycling

Multiple and diverse Proteobacterial species harbored the dissimilatory sulfite reduction enzyme EC:1.8.99.5. These included many putative sulfate-reducing Deltaproteobacteria, several sulfur oxidizing bacteria from the family Sedimenticolaceae, a purple sulfur bacteria ASV, and a purple nonsulfur bacteria ASV.

The sulfate-reducing Deltaproteobacteria fell into four families, with biogeographical patterns more or less conserved within the families. Desulfarculaceae were primarily associated with the smallest size class of particles, though also found on larger particles at all stations south of and including CB3.3C. Desulfobaceraceae had a similar pattern but were sparse at the southernmost station CB5.5. Most members of this phyla were also present on the largest particle size fraction at station CB3.3 and intermediate size fraction at station CB4.3. Desulfovibronaceae appeared to be both free-living and associated with the smallest size fraction (1.2 – 5 μm), though at station 3.3 some were also associated with the largest ≥ 500 μm size class. Desulfobulbaceae at station CB3.3 was most abundant in intermediate size fractions (20 – 53 μm and 53 – 180 μm). A single ASV (Desulfofustis; 617) was also found on intermediate sized particles at the northernmost stations CB3.1 and CB3.2. At stations CB4.3C and CB5.1, this Desulfobaceae was predominantly found on smaller particles. Thus Desulfobulbaceae tended to associate with different particle sizes depending on salinity. Some Desulfobulbaceae can reverse their sulfur reduction pathway (Trojan et al. 2016), and so may oxidize sulfur.

Sulfur oxidizing bacteria included both photosynthetic (purple sulfur and non-sulfur bacteria Bryant and Frigaard 2006) and non-photosynthetic (Sedimenticolaceae) members. Sulfur oxidizing bacteria from the Sedimenticolaceae family appeared to largely co-occur with sulfate reducing bacteria. They were found both free-living and in all size classes, though which size classes they were found in varied between stations. One ASV of purple sulfur bacteria, which was identified to the Chromatiales order appeared primarily in the smallest size class at the most anoxic stations (CB3.3C, CB4.3C and CB5.1), though it also showed up in larger particles at stations CB3.2 and CB3.3C. Similarly, one ASV of purple nonsulfur bacteria, identified to the Rhodospirillaceae family, was found at all Station CB3.3C and south, and was primarily free-living, though at stations CB4.3C and CB5.1 was also associated with the smallest particle size class.

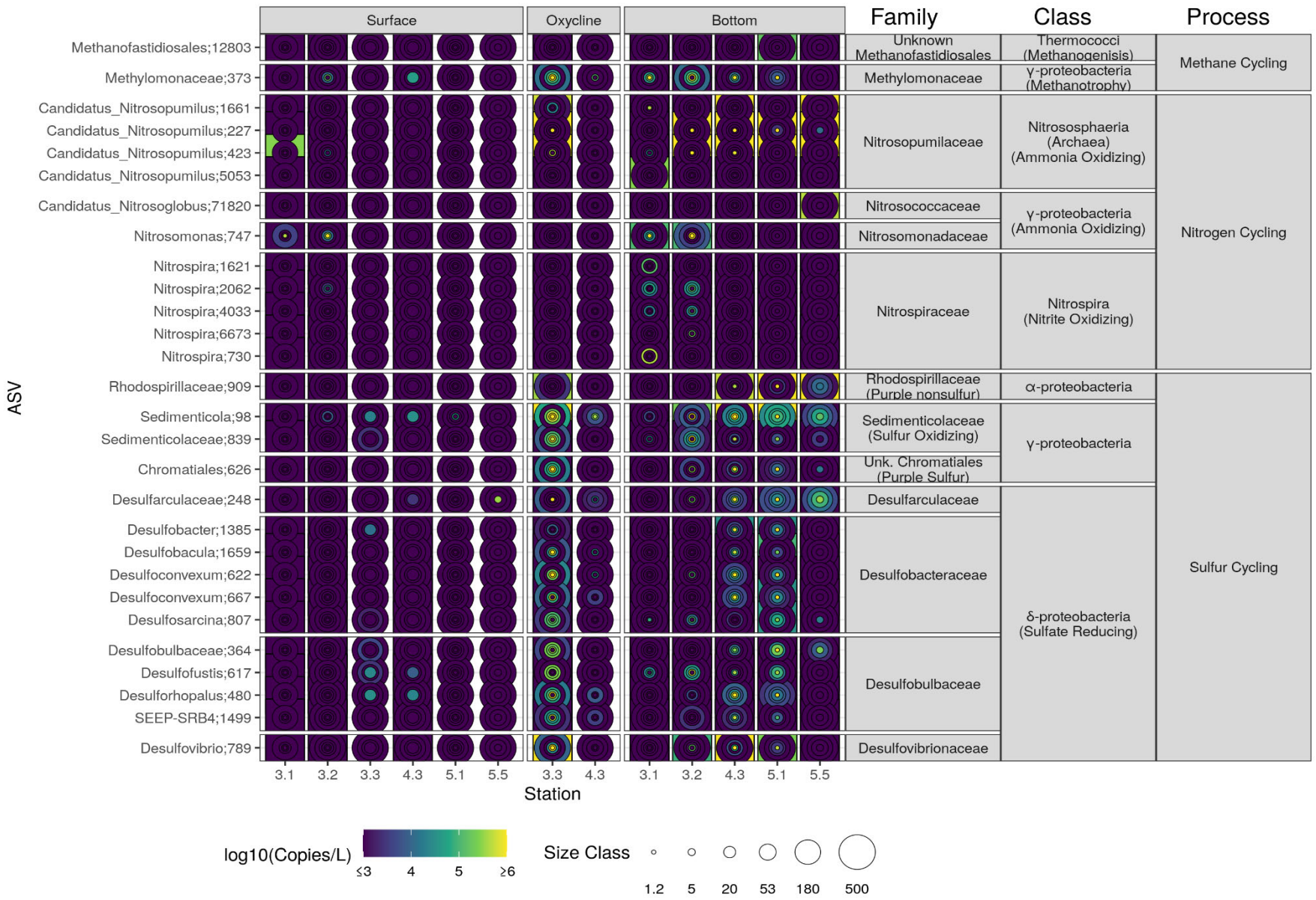

Surface
Oxycline
Bottom
Family
Class
Process
ASV
Methanofastidiosales;12803
Methylomonaceae;373
Candidatus_Nitrosopumilus;1661
Candidatus_Nitrosopumilus;227
Candidatus_Nitrosopumilus;423
Candidatus_Nitrosopumilus;5053
Candidatus_Nitrosoglobus;71820
Nitrosomonas;747
Nitrospira;1621
Nitrospira;2062
Nitrospira;4033
Nitrospira;6673
Nitrospira;730
Rhodospirillaceae;909
Sedimenticola;98
Sedimenticolaceae;839
Chromatiales;626
Desulfarculaceae;248
Desulfobacter;1385
Desulfobacula;1659
Desulfoconvexum;622
Desulfoconvexum;667
Desulfosarcina;807
Desulfobulbaceae;364
Desulfofustis;617
Desulforhopalus;480
SEEP-SRB4;1499
Desulfovibrio;789
Unknown Methanofastidiosales
Thermococci (Methanogenisis)
Methane Cycling
Methylomonaceae
γ-proteobacteria (Methanotrophy)
Nitrosopumilaceae
Nitrososphaeria (Archaea) (Ammonia Oxidizing)
Nitrosococcaceae
γ-proteobacteria (Ammonia Oxidizing)
Nitrosomonadaceae
Nitrospiraceae
Nitrospira (Nitrite Oxidizing)
Nitrogen Cycling
Rhodospirillaceae (Purple nonsulfur)
α-proteobacteria
Sedimenticolaceae (Sulfur Oxidizing)
γ-proteobacteria
Unk. Chromatiales (Purple Sulfur)
Desulfarculaceae
Desulfobacteraceae
δ-proteobacteria (Sulfate Reducing)
Desulfobulbaceae
Desulfovibrionaceae
Sulfur Cycling
3.1
3.2
3.3
4.3
5.1
5.5
3.3
4.3
3.1
3.2
4.3
5.1
5.5
Station
log10(Copies/L)
≤3
4
5
≥6
Size Class
1.2
5
20
53
180
500

Figure 7. Abundance of bacterial ASVs that are putatively involved in methanogenesis, methanotrophy, nitrogen cycling or sulfur cycling (see methods for how functionality was determined). Panels are grouped vertically by biogeochemical *process* type, Class and Family. Parentheses indicate biogeochemical processes in which all members of a given clade are involved, and whether a given Class is from the Archaeal domain. ASV number is indicated after a semicolon.

Discussion

While our approach gives consistent results with previous size fractionation based studies (Mestre et al. 2018), that bacterial communities vary with particle size, location and depth (Figures 3, S11, S12), the more quantitative approach employed in this project extends these results by identifying the sizes of particles in which microorganisms primarily reside. We present six novel observations (I- VI). Specifically, we showed that (I) particle associated microbial abundance scales linearly with particle mass, (II) most organisms are free-living or associated with the smallest particle size class, while only a few organisms associated primarily with larger particles, and (III) there were no abundant free-living bacteria that were not also present on particles. We also showed that (IV) microbial richness is generally highest on particles of intermediate size (5 – 180 μm). Our method allows us to describe (V) the distribution patterns of the eukaryotic community and size and particle partitioning of some members. Finally, this approach allowed us to show that (VI) bacteria involved in anoxic processes associated with the transformation of methane, sulfur and nitrogen were associated primarily with particular particle sizes in particular regions and depths of the Bay. In this discussion we first explore some methodological considerations, and then expand on each of these six novel observations.

Six novel observations

(I) Microbial abundance scales linearly with particle mass

The observation that bacterial abundance scales with particle mass and particulate organic carbon mass suggests that bacteria are likely distributed throughout particles, rather than just on their surface, or that particles are fractal in shape such that their effective surface area scales linearly with their volume. Such an observation is consistent with prior microscopy based observations that bacteria are distributed throughout the core of marine particles (Flintrop et al. 2018). The observation that bacterial abundance is lower (by over two orders of magnitude on the 20 – 53 μm size class than on the same size class at other stations) on particles at the northernmost station CB3.1 could reflect that these northern particles differ in their physics and chemistry from those further south in the Bay, such that they support fewer bacteria relative to their mass. Furthermore, at the bottom of the water column (the only depth this sample was measured), the CB3.1 site had the fewest ASVs associated with the largest particle size fraction (breakaway richness estimate of 100 ASVs, vs 1610-2260 ASVs for all 500 μm, bottom water, samples; Figure S9) and a community structure most distinct from smaller particles (Figure 3), suggesting a distinct environment at this site. Just north of our northernmost site CB3.1, the Bay is characterized by an estuarine turbidity maximum (Schubel 1968). This region has high particle loading and more terrestrial particle origins than elsewhere in the Bay (Schubel 1968; Sanford et al. 2001; Malpezzi et al. 2013). In particular, the turbidity maximum traps particles of intermediate sinking speed (Geyer 1993), which could in principle select for particles with

742 elevated mineral ballast content. However, the particles in this region had carbon to mass ratios
743 and C:N ratios that were similar to those seen elsewhere (Figure 2B), suggesting that relevant
744 chemical differences, if they exist, extend beyond the carbon to mass ratio. Despite this lower
745 microbial abundance on particles, estuarine turbidity maximum systems are typified by fast
746 microbial growth rates (Baross et al. 1994; Lee et al. 2012) especially by particle associated
747 bacteria (Crump and Baross 2000), and by high particle concentrations (Schubel 1968), and so
748 particle associated microbial heterotrophic productivity and other biogeochemical process rates
749 are not necessarily lower at the northernmost site, even though the bacteria are less dense on
750 particles.

751 *(II) Most, but not all, organisms are associated primarily with small particles*

752 We showed, for the first time, that most particle associated organisms are primarily
753 associated with the smallest particles (range 69-99%; Figures 2C, 5, S14). This is true even for
754 taxa that are more abundant relative to particle mass on larger particles, because small particles
755 are so much more abundant by number and mass than large ones (Figure 2A). Such a pattern
756 indicates that the primary habitat of most particle associated bacteria are small particles. As
757 long as microbial growth rates are not orders of magnitude faster on large particles than small
758 ones, this pattern would further suggest that most taxa are adapted to small rather than large
759 particle environments. Small particles are typified by longer residence times than larger particles
760 (Alldredge and Gotschalk 1988; DeVries et al. 2014), which may select for bacteria able to
761 capitalize on these more persistent environments (Kiørboe et al. 2003), and higher spherical
762 surface area to volume ratio allowing for more advection of oxidants throughout the particle
763 (Weber and Bianchi 2020).However, some taxa are primarily associated with intermediate and
764 large particles (Figure 6), including methane and sulfur cycling taxa that may favor anoxic
765 microenvironments (Figure 7).

766 Several ASVs showed exceptions to the above pattern and are primarily associated with
767 particles 20 μm or larger. Some of these bacteria such as Midochloraceae (Montagna et al.
768 2013) and Sulfitobacter (Amin et al. 2015; Johansson et al. 2019; Shibl et al. 2020) are known
769 to be symbionts and may be associated with larger organisms and likely appear in our larger
770 size fractions because their hosts partition into larger size classes. In contrast, the *Paracoccus*
771 ASV which was found to associate statistically with large particles likely associates physically
772 with those particles, rather than large organisms. *Paracoccus* can both grow aerobically and can
773 break down a variety of sugars and other compounds (Harker et al. 1998; Tsubokura et al.
774 1999). Thus, it seems probable given their metabolism that the *Paracoccus* ASV is degrading
775 the large particles. An ASV from the *Vibrio* genus, which we could not classify further with NCBI-
776 BLAST, was associated especially with the largest size class. Some but not all *Vibrio* are
777 pathogens (Colwell et al. 1977) and may be associated with zooplankton (Kaneko and Colwell
778 1973), but many are also are known to associate with suspended particles (Froelich et al. 2013;
779 Kirstein et al. 2016; Liang et al. 2019), and could be in either role in this environment.

780 *(III) We observed no exclusively free-living, abundant, bacteria*

781 The observation that every ASV that comprised ≥ 1% of the free-living community was also
782 found associated with some particle size fraction indicates that all abundant bacteria in the Bay
783 mainstem are at least sometimes associated with particles, at least at the time and locations

where we sampled. This is in contrast to observations of the Baltic Sea, where it was shown that many bacterial species were only found in the free-living fraction (0.22 – 5 μm; Rieck et al. 2015). However, SAR11 bacteria which are believed to be free-living in nature (Giovannoni et al. 2005) and grow free-living in culture (Rappé et al. 2002), are seen not only associated with particles in our dataset (Figure S15), but also in large size fractions in the more oligotrophic Blanes Bay microbial observatory in the Mediterranean Sea (Mestre et al. 2020) or in the Eastern Tropical North and South Pacific (Ganesh et al. 2014; Fuchsman et al. 2017), suggesting that the association of abundant "free-living" bacteria with particles may happen elsewhere. One possibility is that free-living bacteria become associated with particles through physical processes, such as sticking and aggregation, which the bacteria cannot avoid. This has been shown for *Synechoccocus* in the laboratory (Cruz and Neuer 2019). Additionally, viral infection can cause bacteria to clump (Shibata et al. 1997), and high viral loads on particles of viruses that infect free-living bacteria support the importance of this pathway (Ganesh et al. 2014; Fuchsman et al. 2019a). As particles are particularly abundant in the tidal Chesapeake Bay (Dougherty et al. 2021; Turner et al. 2021) perhaps physical aggregation is more pronounced in the Bay than elsewhere.

(IV) Microbial richness is highest on intermediate sized particles

The observation that particles of intermediate size (5 – 180 μm) harbored highest richness, and thus highest Shannon diversity could indicate that intermediate sized particles have characteristics of both larger and smaller particles and so harbor communities typical of both particle types. In other ocean sites, high richness has been seen in transitions between different communities. For instance, in the oligotrophic coastal ocean richness was high in the mesopelagic transition between the surface deep ocean environments (Cram et al. 2015). In an estuary system, richness was shown to be highest in brackish water (Tee et al. 2021). Meanwhile in sediments, nitrifying bacteria appear to be most diverse at zones of redox transitions (Zhao et al. 2019).. Perhaps a similar pattern happens along the observed particle size gradient with intermediate sized particles containing attributes and microorganisms from both larger and smaller size fractions.

The lower diversity in the central Bay opposes patterns seen in the Columbia River and Waiwera River estuaries, where alpha diversity was higher in brackish waters than elsewhere (Fortunato et al. 2011; Tee et al. 2021). In our system, the sulfidic waters in the brackish section of the estuary likely select against many common bacteria. This brackish and sulfidic bottom environment, because it is smaller than the oceans or watershed systems that surround it, may be affected by island biogeography effects (MacArthur and Wilson 2001) in which smaller systems support fewer species.

(V) Algal and zooplankton size and spatial distribution patterns

In contrast to bacterial phyla, most Eukaryotic phyla appear to have patchy distribution across space and are often found only associated with particular size classes (Figure 4). Ochrophyta (diatoms and other brown algae) and Ciliophora (cilliates) are microorganisms that were abundant in our largest size classes, suggesting that they may aggregate into and/or associate with particles. (Figure 4).

The observation that zooplankton associate with particular size classes reflects the size of those organisms. For instance, Arthropods, dominated by Maxilopods, were found primarily in the 53-180 μm size class, suggesting that we have primarily sampled individuals from the Nauplii or Copepodite life stages. The dominance of sub-adult life stages is consistent with previous observations (Kimmel et al. 2006), and the fact that adult maxilopods can avoid being collected in Niskin bottles. Alternatively, these findings could reflect fragments of zooplankton carapaces and molts being collected on the smaller size filters. The observation that arthropods are mainly in the Southernmost station, and that they avoid the sulfidic region of the Bay also reflect previous observations (Zhang et al. 2006). Meanwhile Ctenophora are primarily associated with our larger size classes reflecting that ctenophores and parts thereof tend to be larger than 500 μm (Ruppert et al. 2004).

(VI) Elemental Cycling

Bacteria involved in methane, sulfur and nitrogen cycling were shown to have particle size and water salinity specific habitats (Figure 7). Putative methane, sulfur and nitrogen cycling organisms were each most abundant in bottom waters and scarce in surface waters, confirming that anoxia and/or interaction with the sediment are likely important for all of these processes. Sulfate reducing and sulfide oxidizing microorganisms have been associated with particles in other anoxic systems (Fuchsman et al. 2012, 2017; Saunders et al. 2019; Raven et al. 2021), but their association with particular particle sizes has not been seen before, to our knowledge.

Methane cycle

The observation that methanogens were scarce in our dataset, while methanotrophs were abundant on particles supports the inference that methanogenesis likely occurs in the sediment (as shown by Gelesh et al. 2016) and not on suspended particles; but methane is consumed in the water column primarily by particle associated, rather than free-living, bacteria. Methanotroph ASVs were present at all stations except station CB5.5 (Figure 7), suggesting that methane is produced by the sediment across the Bay and consumed in the overlying waters. CB5.5 is more marine, and this region has lower organic carbon in its sediments (Roden and Tuttle 1993; Zimmerman and Canuel 2001). Thus, methane is either not produced, or is both produced and consumed in the sediments.

Nitrogen cycle

Our data suggest that ammonia-oxidizing bacteria and archaea are ubiquitous in the bottom waters of the Chesapeake Bay, and are primarily associated with the free-living (0.2 – 1.2 μm) and next-smallest (1.2 – 5 μm) size fractions. The presence of these organisms in the anoxic bottom waters of the Bay is surprising, since these nitrifying bacteria and archaea require oxygen as an oxidant. The ammonia-oxidizing organisms could be either advected or dispersed into the anoxic bottom waters. However, distinct ammonia-oxidizing archaea have also been found in the upper sulfidic zone of the Black Sea, implying a more complex lifestyle (Coolen et al. 2007). Large particles appear not to be important habitats for ammonium oxidizing organisms. The exception to this pattern was an ASV from the *Nitrosomonas* genus, which appeared to associate with larger particles in the northernmost stations of the Bay (Figure 7). *Nitrosomonas* has been found associated with particles previously in the Mediterranean Sea (Phillips et al. 1999).

In contrast, nitrite-oxidizing ASVs from the *Nitrospira* genus appeared to have, as their primary habitat, particles from the northernmost Bay stations. Such a pattern could suggest geographic decoupling between ammonium and nitrite oxidation. However, the lack of observed nitrite oxidizers in the southern station despite the abundance of ammonium oxidizing archaea suggests that some other unidentified organisms are likely consuming the nitrite produced by the ammonium oxidizers at these stations.

Our PCR primers have mismatches to anammox bacteria (McNichol et al. 2021), and unsurprisingly we did not detect any anammox bacteria in our dataset. Future efforts should consider using the pooled primer sets that have since been described by McNichol et al. (2021), and which better amplify organisms from this group. We did not look for denitrifying bacteria in our site, as denitrification is not phylogenetically conserved (Zumft 1997; Bertagnolli et al. 2020) and so neither taxonomic groupings nor phylogenetic placement is likely to generate reliable information about this process. However this process is common in anoxic systems, and has been measured in the Chesapeake previously (Ji et al. 2018). Furthermore, new evidence suggests that many of the taxa associated with sulfur cycling in the Chesapeake also reduce nitrate (Arora-Williams et al. 2022). Thus, it is likely that many bacteria especially in the anoxic waters are indeed removing nitrogen from the system.

Sulfur cycle

Sulfur cycling organisms were particularly abundant at the oxycline sample of station CB3.3C (Figure 7). These sulfur cycling taxa were not as abundant at station CB4.3C's oxycline depth. Consistent with this, our sulfide profiles indicate that CB3.3 had higher sulfide concentrations than CB4.3 (Figure S3). These differences could be due to variations in the benthic sulfide flux (Roden and Tuttle 1992), which can play a role in driving sulfur cycling in anoxic bottom waters. While sulfide transport out of sediment occurs in the summer at both sites, site CB3.3C has higher sulfate reduction rates in the winter which may promote higher porewater sulfide concentrations and sulfide transport out of sediment at this site (Roden and Tuttle 1993). Sulfur cycling organisms, especially sulfate reducing Deltaproteobacteria, and the sulfur oxidizing and purple sulfur bacteria from the Gammaproteobacteria, were found associated with large particles. The exception was purple nonsulfur bacteria which were found to be free-living. Purple sulfur bacteria are large, with a width of around 4 μm (Madigan and Martinko 2005), and so their presence in the 1.2 – 5 μm size fraction could indicate that these organisms were free-living when found in that size fraction. However, these purple sulfur bacteria were also found in the larger size factions suggesting they were also particle associated. Several ASVs of sulfate reducing Desulfobacteraceae and Desulfobaeae were associated with intermediate particle size fractions at station CB3.2, rather than the smallest particle size fraction. This pattern is unusual since most ASVs are primarily on the small size fractions, as discussed above. Therefore, this abundance on intermediate particles suggests that these bacteria, at least at station CB3.2, have intermediate particles as their primary habitat. It has been suggested that large particles may be sites of sulfate reduction because reductants are abundant and other oxidants such as nitrate, nitrite and oxygen cannot diffuse into the particle cores (Bianchi et al. 2018) and sulfate reducers and sulfate reduction have been found on particles previously (Fuchsman et al. 2012; Saunders et al. 2019; Raven et al. 2021). Such a pattern then suggests the presence of anoxic environments in these large particles

which may be leveraged by sulfate reducing bacteria. However, the association of sulfate reducers primarily with the smallest particle size fraction at the other stations suggests a more complex pattern.

Methodological Considerations

Our method worked well to quantitatively measure microbial community structure along a particle size gradient, across the surface and deep waters of the Chesapeake Bay. However, there are some important methodological considerations intrinsic to this approach. Most of these are not unique to our study, but rather affect microbial community studies in general, but warrant consideration here. The considerations are that some bacteria are large, some particles are fragile, some bacteria contain multiple copies of marker genes, and that Illumina read quality was comparatively lower for this project than in other amplicon studies. *Large Bacteria:* Some bacterial taxa exceed 1.2 µm in width and so may show up in the 1.2 – 5 µm bin. However, our primary conclusions hold even if we only consider bacteria in the 5 µm and larger (Figure S12A,B). *Fragile Particles:* The delicate nature of particles introduces the risk of disaggregation during gravity filtration, potentially skewing the particle size to mass spectrum and reassigning cells to smaller size fractions. *Filter Handling:* Despite precautions, including frequent rinsing of filters, there is a potential for filter clogging and bacterial aggregation during filtration, especially for the smaller mesh sizes, potentially leading to misrepresentation of bacterial abundance and community structure in some size categories. *Genetic markers:* Many bacteria contain multiple (between 1 and 15) copies of the 16S gene (Venter et al. 2004; Espejo and Plaza 2018) and eukaryotes even more copies of the 18S gene (Fuchsman et al. 2022 and references therein), which inflate gene copy counts and skew proportions towards bacteria with more 16S gene copies. Microscopy validation indicates that our DNA-based methods are realistic, though higher on average than microscopy counts (Figure S8). *DNA Extraction Efficiency*: Our assumption of 100% DNA extraction efficiency (Eqn. 2). may lead to underestimates of the abundance of some taxa (Han et al. 2018; Nearing et al. 2021; Brauer and Bengtsson 2022). Adding spikes before DNA extraction (following Gifford et al. 2020) could correct for this loss, but was not done here. *Amplification Bias:* The PCR step of amplicon sequencing may preferentially amplify some organisms (Elbrecht and Leese 2015). While our primer set has been validated and shows minimal bias, it does not capture all bacterial clades (McNichol et al. 2021). *Read Quality Issues:* Our lower than typical read quality (Lee 2019) might have adversely impacted accuracy (see Results, Amplicon Processing and Figure S7A-D). However, similarities in richness between different sequencing runs affirms that read quality does not substantially affect our results. Despite this, the presence of low read number ASVs could artificially increase perceived diversity. *Low Read Depth Samples:* We included three samples with low read depth (< 1300 reads, after processing), as they provided information about community structure. While they may have introduced inaccuracies in some of our ordination analyses, we used diversity index statistics that were robust to the low read depths.

Additional points of discussion

Samples of free-living and small particle associated microorganisms appear to vary more with temperature and salinity than microorganisms that are associated with larger particles (Figure 3). One possible reason for this greater similarity between large particle samples is that microbial niches on large particles might be defined more by the properties of the particle to

which they are attached than to the surrounding water. Meanwhile free-living bacteria have the water itself as their habitat and small particle-attached bacteria may experience conditions intermediate to those of free-living bacteria and those attached to large particles – for instance they likely have some of the chemical substrates and physical attachment surfaces in common with large particle-attached bacteria, but exposure to oxygen or other oxidants more like those of free-living bacteria. Previous analysis has shown similarity of microbes on particles across depths in the oligotrophic ocean and much of that was attributed to particle movement between environments (Mestre et al. 2018), which could also be a factor here.

While we found that microbial abundance scales with particle mass, we did not try to normalize to particle surface area. This is because marine snow are believed to have shapes that are closer to fractals than spheres, with substantial folding and pore space (Logan and Wilkinson 1990; Dissanayake et al. 2018). Thus, they have surface areas much higher than for equivalently sized spheres. Furthermore, bacteria are found throughout particles (Flintrop et al. 2018). Thus, we contend that it is reasonable that microbial abundance scales with particle volume, which, assuming that the volume of the non-porous portion of the particle scales with mass (as in Cram et al. 2018), means that microbial abundance should also scale with mass.

Conclusion

This work extends our knowledge of microbial biogeography in the Chesapeake Bay from an analysis of spatial and depth variability into one that concurrently considers variability of habitats within locations. We show that such within-location habitat variability is important and that bacteria have niches that are defined not only by geography, but also particle size. Bacteria involved in sulfur and methane cycling appear to associate with intermediate particle size classes, suggesting that select particles provide sites for water column biogeochemical transformations that often occur in anoxic environments. Extending this approach to consider bacterial biogeography within habitats in more regions, and across more kinds of microhabitats will provide new understanding of the ecological niches of marine microorganisms.

Data Availability Statement

Sequence data are available on NCBI's Short read archive under accession number PRJNA898904 "Chesapeake Bay 2019 Particle Size Fractionation." The amplicon sequencing pipeline, applied to the raw sequence data and which generates the microbial read counts and taxonomy is archived and publicly available on FigShare https://doi.org/10.6084/m9.figshare.21950354.v1. Data analysis scripts used to generate all figures and tables are also on FigShare https://doi.org/10.6084/m9.figshare.21948425.v2.

Acknowledgements

The authors thank Sairah Malkin, Mekayla Reynolds, Emily Dougherty and the Captain and crew of the UMCES RV *Rachael Carson* for their help with field collection; Sabeena Nazar and Allan Place for their assistance with sequencing; Michael Lee and Chris Trivedi for guidance in amplicon sequence data processing; and Larry Sanford for useful feedback regarding methodology. Ryan Moore provided insight about alpha diversity measurement. Jeff Cornwell and Michael Owens produced the hydrogen sulfide profile at station CB4.3C. James Pierson and Catherine Fitzgerald provided helpful insight about the taxonomy ecology and development of mesoplankton, especially *Acartia tonsa*. João Maia designed the graphical abstract. The

authors are thankful for the Startup funds provided to Cram and Fuchsman by UMCES Horn Point Laboratory. G. Martinez was funded by the Maryland Sea Grant REU program which in turn was funded by NSF grant OCE-1756244.

References

Alldredge AL, Cole JJ, Caron DA. Production of heterotrophic bacteria inhabiting macroscopic organic aggregates(marine snow) from surface waters. Limnol Oceanogr. 1986;31(1):68–78.

Alldredge AL, Gotschalk C. In situ settling behavior of marine snow. Limnol Oceanogr. 1988 May 1;33(3):339–51.

Alldredge AL, Silver MW. Characteristics, dynamics and significance of marine snow. Prog Oceanogr. 1988;20(1):41–82.

Amin SA, Hmelo LR, van Tol HM, Durham BP, Carlson LT, Heal KR, et al. Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. Nature. 2015 Jun;522(7554):98–101.

Arora-Williams K, Holder C, Secor M, Ellis H, Xia M, Gnanadesikan A, et al. Abundant and persistent sulfur-oxidizing microbial populations are responsive to hypoxia in the Chesapeake Bay. Environ Microbiol [Internet]. 2022 [cited 2022 Apr 6];00(00). Available from: http://onlinelibrary.wiley.com/doi/abs/10.1111/1462-2920.15976

Arthur MA, Dean WE, Claypool GE. Anomalous 13C enrichment in modern marine organic carbon. Nature. 1985 May;315(6016):216–8.

Azam F, Malfatti F. Microbial structuring of marine ecosystems. Nat Rev Microbiol. 2007 Oct;5(10):782–91.

Barbera P, Kozlov AM, Czech L, Morel B, Darriba D, Flouri T, et al. EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. Syst Biol. 2019 Mar 1;68(2):365–9.

Baross J, Crump B, Simenstad C. Elevated ‘microbial loop’ activities in the Columbia River estuary turbidity maximum. In: Changing Particle Fluxes in Estuaries: Implications from Science to Management. 1994. p. 459–64.

Bertagnolli AD, Konstantinidis KT, Stewart FJ. Non-denitrifier nitrous oxide reductases dominate marine biomes. Environ Microbiol Rep. 2020;12(6):681–92.

Bianchi D, Weber TS, Kiko R, Deutsch C. Global niche of marine anaerobic metabolisms expanded by particle microenvironments. Nat Geosci. 2018 Apr;11(4):263.

Bidle KD, Fletcher M. Comparison of free-living and particle-associated bacterial communities in the chesapeake bay by stable low-molecular-weight RNA analysis. Appl Environ Microbiol. 1995;61(3):944–52.

Boyd CE. Hydrogen Sulfide Toxic, But Manageable. Glob Aquac Advocate. 2014;34–6.

Brauer A, Bengtsson MM. DNA extraction bias is more pronounced for microbial eukaryotes than for prokaryotes. MicrobiologyOpen. 2022;11(5):e1323.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016 Jul;13(7):581–3.

Colwell RR, Kaper J, Joseph SW. Vibrio cholerae, Vibrio parahaemolyticus, and Other Vibrios: Occurrence and Distribution in Chesapeake Bay. Science. 1977 Oct 28;198(4315):394–6.

Coolen MJL, Abbas B, Van Bleijswijk J, Hopmans EC, Kuypers MMM, Wakeham SG, et al. Putative ammonia-oxidizing Crenarchaeota in suboxic waters of the Black Sea: a basin-wide ecological study using 16S ribosomal and functional genes and membrane lipids. Environ Microbiol. 2007;9(4):1001–16.

Cram JA, Chow CET, Sachdeva R, Needham DM, Parada AE, Steele JA, et al. Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. ISME J. 2015;9:563–80.

Cram JA, Parada AE, Fuhrman JA. Dilution reveals how viral lysis and grazing shape microbial communities: Viral Lysis and Grazing Shape Microbial Communities. Limnol Oceanogr. 2016 Feb;61(3):889–905.

Cram JA, Weber T, Leung SW, McDonnell AMP, Liang JH, Deutsch C. The Role of Particle Size, Ballast, Temperature, and Oxygen in the Sinking Flux to the Deep Sea. Glob Biogeochem Cycles. 2018 May;32(5):858–76.

Crump BC, Peranteau C, Beckingham B, Cornwell JC. Respiratory Succession and Community Succession of Bacterioplankton in Seasonally Anoxic Estuarine Waters. Appl Environ Microbiol. 2007 Nov 1;73(21):6802–10.

Cruz BN, Neuer S. Heterotrophic Bacteria Enhance the Aggregation of the Marine Picocyanobacteria Prochlorococcus and Synechococcus. Front Microbiol. 2019 Aug 13;10:1864.

Czech L, Barbera P, Stamatakis A. Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. Bioinformatics. 2020 May 1;36(10):3263–5.

DeLong EF, Franks DG, Alldredge AL. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. Limnol Oceanogr. 1993;38(5):924–34.

DeVries T, Liang JH, Deutsch C. A mechanistic particle flux model applied to the oceanic phosphorus cycle. Biogeosciences Discuss. 2014 Mar 6;11(3):3653–99.

Dissanayake AL, Burd AB, Daly KL, Francis S, Passow U. Numerical Modeling of the Interactions of Oil, Marine Snow, and Riverine Sediments in the Ocean. J Geophys Res Oceans. 2018;123(8):5388–405.

Dougherty E, Cram J, Hollins A. The relationship between size, abundance, and mass of particles in the surface and bottom waters of the Chesapeake Bay [Internet]. Oceanography; 2021 Oct. Available from: http://www.essoar.org/doi/10.1002/essoar.10508364.1

1070 Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2 for
1071 prediction of metagenome functions. Nat Biotechnol. 2020 Jun;38(6):685–8.

1072 Elbrecht V, Leese F. Can DNA-Based Ecosystem Assessments Quantify Species Abundance?
1073 Testing Primer Bias and Biomass—Sequence Relationships with an Innovative
1074 Metabarcoding Protocol. Hajibabaei M, editor. PLOS ONE. 2015 Jul 8;10(7):e0130324.

1075 Espejo RT, Plaza N. Multiple Ribosomal RNA Operons in Bacteria; Their Concerted Evolution
1076 and Potential Consequences on the Rate of Evolution of Their 16S rRNA. Front
1077 Microbiol. 2018;9:1232.

1078 Farnelid H, Turk-Kubo K, Ploug H, Ossolinski JE, Collins JR, Mooy BASV, et al. Diverse
1079 diazotrophs are present on sinking particles in the North Pacific Subtropical Gyre. ISME
1080 J. 2018 Aug 16;1.

1081 Findlay AJ, Bennett AJ, Hanson TE, Luther GW. Light-Dependent Sulfide Oxidation in the
1082 Anoxic Zone of the Chesapeake Bay Can Be Explained by Small Populations of
1083 Phototrophic Bacteria. Appl Environ Microbiol. 2015 Nov;81(21):7560–9.

1084 Flintrop CM, Rogge A, Miksch S, Thiele S, Waite AM, Iversen MH. Embedding and slicing of
1085 intact in situ collected marine snow: Embedding and slicing of marine snow. Limnol
1086 Oceanogr Methods. 2018 Jun;16(6):339–55.

1087 Fortunato CS, Herfort L, Zuber P, Baptista AM, Crump BC. Spatial variability overwhelms
1088 seasonal patterns in bacterioplankton communities across a river to ocean gradient.
1089 ISME J. 2011;6(3):554–63.

1090 Froelich B, Ayrapetyan M, Oliver JD. Integration of Vibrio vulnificus into Marine Aggregates and
1091 Its Subsequent Uptake by Crassostrea virginica Oysters. Appl Environ Microbiol. 2013
1092 Mar;79(5):1454–8.

1093 Fuchsman CA, Cherubini L, Hays MD. An analysis of protists in Pacific oxygen deficient zones:
1094 implications for *Prochlorococcus* and N_2 -producing bacteria. Environ Microbiol. 2022
1095 Apr;24(4):1790–804.

1096 Fuchsman CA, Devol AH, Saunders JK, McKay C, Rocap G. Niche Partitioning of the N Cycling
1097 Microbial Community of an Offshore Oxygen Deficient Zone. Front Microbiol [Internet].
1098 2017 Dec 5 [cited 2018 Dec 4];8. Available from:
1099 http://journal.frontiersin.org/article/10.3389/fmicb.2017.02384/full

1100 Fuchsman CA, Murray JW, Staley JT. Stimulation of Autotrophic Denitrification by Intrusions of
1101 the Bosporus Plume into the Anoxic Black Sea. Front Microbiol [Internet]. 2012 [cited
1102 2019 Jul 6];3. Available from:
1103 https://www.frontiersin.org/articles/10.3389/fmicb.2012.00257/full

1104 Fuchsman CA, Palevsky HI, Widner B, Duffy M, Carlson MCG, Neibauer JA, et al.
1105 Cyanobacteria and cyanophage contributions to carbon and nitrogen cycling in an
1106 oligotrophic oxygen-deficient zone. ISME J. 2019a Jun 27;1.

Fuchsman CA, Paul B, Staley JT, Yakushev EV, Murray JW. Detection of Transient Denitrification During a High Organic Matter Event in the Black Sea. Glob Biogeochem Cycles. 2019b;33(2):143–62.

Fuhrman JA, Comeau DE, Hagström \AAke, Chan AM. Extraction from natural planktonic microorganisms of DNA suitable for molecular biological studies. Appl Environ Microbiol. 1988;54(6):1426–9.

Ganesh S, Parris DJ, DeLong EF, Stewart FJ. Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. ISME J. 2014 Jan;8(1):187–211.

Garcia JL, Patel BKC, Ollivier B. Taxonomic, Phylogenetic, and Ecological Diversity of Methanogenic Archaea. Anaerobe. 2000 Aug 1;6(4):205–26.

Gelesh L, Marshall K, Boicourt W, Lapham L. Methane concentrations increase in bottom waters during summertime anoxia in the highly eutrophic estuary, Chesapeake Bay, U.S.A. Limnol Oceanogr. 2016;61(S1):S253–66.

Geyer WR. The Importance of Suppression of Turbulence by Stratification on the Estuarine Turbidity Maximum. Estuaries. 1993 Mar;16(1):113.

Gifford SM, Zhao L, Stemple B, DeLong K, Medeiros PM, Seim H, et al. Microbial Niche Diversification in the Galápagos Archipelago and Its Response to El Niño. Front Microbiol [Internet]. 2020 [cited 2022 Jul 11];11. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2020.575194

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, et al. Genome Streamlining in a Cosmopolitan Oceanic Bacterium. Science. 2005 Aug 19;309(5738):1242–5.

Hagen RA, Vogt PR. Seasonal variability of shallow biogenic gas in Chesapeake Bay. Mar Geol. 1999 Jun 1;158(1):75–88.

Han Z, Sun J, Lv A, Wang A. Biases from different DNA extraction methods in intestine microbiome research based on 16S rDNA sequencing: a case in the koi carp, Cyprinus carpio var. Koi. MicrobiologyOpen. 2018 Apr 17;8(1):e00626.

Harker M, Hirschberg J, Oren A 1998. Paracoccus marcusii sp. nov., an orange Gram-negative coccus. Int J Syst Evol Microbiol. 1998;48(2):543–8.

Ji Q, Frey C, Sun X, Jackson M, Lee YS, Jayakumar A, et al. Nitrogen and oxygen availabilities control water column nitrous oxide production during seasonal anoxia in the Chesapeake Bay. Biogeosciences. 2018 Oct 18;15(20):6127–38.

Johansson ON, Pinder MIM, Ohlsson F, Egardt J, Töpel M, Clarke AK. Friends With Benefits: Exploring the Phycosphere of the Marine Diatom Skeletonema marinoi. Front Microbiol [Internet]. 2019 [cited 2022 Oct 25];10. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01828

Kan J, Crump BC, Wang K, Chen F. Bacterioplankton community in Chesapeake Bay: Predictable or random assemblages. Limnol Oceanogr. 2006 Sep;51(5):2157–69.

1144 Kan J, Suzuki MT, Wang K, Evans SE, Chen F. High Temporal but Low Spatial Heterogeneity
1145 of Bacterioplankton in the Chesapeake Bay. Appl Environ Microbiol. 2007 Nov
1146 1;73(21):6776–89.

1147 Kanehisa M. Enzyme Annotation and Metabolic Reconstruction Using KEGG. Methods Mol Biol
1148 Clifton NJ. 2017;1611:135–45.

1149 Kaneko T, Colwell RR. Ecology of Vibrio parahaemolyticus in Chesapeake Bay. J Bacteriol.
1150 1973 Jan;113(1):24–32.

1151 Kang JC. Acute toxicity of hydrogen sulfide to larvae and adults of blue crab Portunus
1152 trituberculatus white shrimp Metapenaeus monoceros and prawn Macrobrachium
1153 nipponens. J Fish Pathol. 1997;10(1):65–72.

1154 Kimmel DG, Roman MR, Zhang X. Spatial and temporal variability in factors affecting
1155 mesozooplankton dynamics in Chesapeake Bay: Evidence from biomass size spectra.
1156 Limnol Oceanogr. 2006 Jan;51(1):131–41.

1157 Kiørboe T, Tang K, Grossart HP, Ploug H. Dynamics of Microbial Communities on Marine Snow
1158 Aggregates: Colonization, Growth, Detachment, and Grazing Mortality of Attached
1159 Bacteria. Appl Environ Microbiol. 2003 Jun 1;69(6):3036–47.

1160 Kirstein IV, Kirmizi S, Wichels A, Garin-Fernandez A, Erler R, Löder M, et al. Dangerous
1161 hitchhikers? Evidence for potentially pathogenic Vibrio spp. on microplastic particles.
1162 Mar Environ Res. 2016 Sep 1;120:1–8.

1163 Laperriere SM, Nidzieko NJ, Fox RJ, Fisher AW, Santoro AE. Observations of Variable
1164 Ammonia Oxidation and Nitrous Oxide Flux in a Eutrophic Estuary. Estuaries Coasts.
1165 2019 Jan 1;42(1):33–44.

1166 Lee D, Keller D, Crump B, Hood R. Community metabolism and energy transfer in the
1167 Chesapeake Bay estuarine turbidity maximum. Mar Ecol Prog Ser. 2012 Mar 8;449:65–
1168 82.

1169 Lee M. Happy Belly Bioinformatics: an open-source resource dedicated to helping biologists
1170 utilize bioinformatics. J Open Source Educ. 2019 Sep 29;2(19):53.

1171 Leu AO, Eppley JM, Burger A, DeLong EF. Diverse Genomic Traits Differentiate Sinking-
1172 Particle-Associated versus Free-Living Microbes throughout the Oligotrophic Open
1173 Ocean Water Column. Bello MGD, editor. mBio. 2022 Jul 12;e01569-22.

1174 Liang J, Liu J, Wang X, Lin H, Liu J, Zhou S, et al. Spatiotemporal Dynamics of Free-Living and
1175 Particle-Associated Vibrio Communities in the Northern Chinese Marginal Seas. Appl
1176 Environ Microbiol. 2019 Apr 18;85(9):e00217-19.

1177 Logan BE, Wilkinson DB. Fractal geometry of marine snow and other biological aggregates.
1178 Limnol Oceanogr. 1990;35(1):130–6.

1179 Long RA, Azam F. Microscale patchiness of bacterioplankton assemblage richness in seawater.
1180 Aquat Microb Ecol. 2001 Dec 5;26(2):103–13.

1181 Louca S, Doebeli M. Efficient comparative phylogenetics on large trees. Bioinformatics. 2018
1182 Mar 15;34(6):1053–5.

1183 MacArthur RH, Wilson EO. The theory of island biogeography. Princeton: Princeton University
1184 Press; 2001.

1185 Madigan MT, Martinko JM. Brock Biology Of Microorganisms 11th EDITION. 11th ed. Lebanon,
1186 Indiana, U.S.A.: Pren; 2005.

1187 Malpezzi MA, Sanford LP, Crump BC. Abundance and distribution of transparent exopolymer
1188 particles in the estuarine turbidity maximum of Chesapeake Bay. Mar Ecol Prog Ser.
1189 2013;486:23–35.

1190 McNichol J, Berube PM, Biller SJ, Fuhrman JA. Evaluating and Improving Small Subunit rRNA
1191 PCR Primer Coverage for Bacteria, Archaea, and Eukaryotes Using Metagenomes from
1192 Global Ocean Surveys. mSystems. 2021 Jun 29;6(3):e0056521.

1193 Mestre M, Borrull E, Sala Mm, Gasol JM. Patterns of bacterial diversity in the marine planktonic
1194 particulate matter continuum. ISME J. 2017 Apr;11(4):999–1010.

1195 Mestre M, Höfer J, Sala MM, Gasol JM. Seasonal Variation of Bacterial Diversity Along the
1196 Marine Particulate Matter Continuum. Front Microbiol [Internet]. 2020 [cited 2022 Jun
1197 2];11. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2020.01590

1198 Mestre M, Ruiz-González C, Logares R, Duarte CM, Gasol JM, Sala MM. Sinking particles
1199 promote vertical connectivity in the ocean microbiome. Proc Natl Acad Sci. 2018 Jun
1200 27;201802470.

1201 Montagna M, Sassera D, Epis S, Bazzocchi C, Vannini C, Lo N, et al. “Candidatus
1202 Midichloriaceae” fam. nov. (Rickettsiales), an Ecologically Widespread Clade of
1203 Intracellular Alphaproteobacteria. Appl Environ Microbiol. 2013 May;79(10):3241–8.

1204 Nearing JT, Comeau AM, Langille MGI. Identifying biases and their potential solutions in human
1205 microbiome studies. Microbiome. 2021 May 18;9(1):113.

1206 Needham D, Fichot E, Parada A, Yeh YC, Fuhrman J. Fuhrman Lab 515F-926R16S and 18S
1207 rRNA Gene Sequencing Protocol v2 [Internet]. 2018 Nov. Available from:
1208 https://www.protocols.io/view/fuhrman-lab-515f-926r-16s-and-18s-rrna-gene-sequen-
1209 vb7e2rn

1210 Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O’Hara RB, et al. vegan:
1211 Community Ecology Package [Internet]. 2013. Available from: http://CRAN.R-
1212 project.org/package=vegan

1213 Palinkas CM, Testa JM, Cornwell JC, Li M, Sanford LP. Influences of a River Dam on Delivery
1214 and Fate of Sediments and Particulate Nutrients to the Adjacent Estuary: Case Study of
1215 Conowingo Dam and Chesapeake Bay. Estuaries Coasts. 2019 Dec 1;42(8):2072–95.

1216 Phillips CJ, Smith Z, Embley TM, Prosser JI. Phylogenetic Differences between Particle-
1217 Associated and Planktonic Ammonia-Oxidizing Bacteria of the β Subdivision of the Class

Proteobacteria in the Northwestern Mediterranean Sea. Appl Environ Microbiol. 1999 Feb;65(2):779–86.

Ploug H, Kühl M, Buchholz B, Jørgensen BB. Anoxic aggregates an ephemeral phenomenon in the ocean. Aquat Microb Ecol. 1997;13:285–94.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013 Jan 1;41(D1):D590–6.

Rappé MS, Connon SA, Vergin KL, Giovannoni SJ. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature. 2002 Aug;418(6898):630–3.

Raven MR, Keil RG, Webb SM. Microbial sulfate reduction and organic sulfur formation in sinking marine particles. Science. 2021 Jan 8;371(6525):178–81.

Reeburgh WS. Observations of Gases in Chesapeake Bay Sediments1. Limnol Oceanogr. 1969;14(3):368–75.

Rieck A, Herlemann DPR, Jürgens K, Grossart HP. Particle-Associated Differ from Free-Living Bacteria in Surface Waters of the Baltic Sea. Front Microbiol [Internet]. 2015 [cited 2022 Nov 8];6. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2015.01297

Roden EE, Tuttle JH. Sulfide release from estuarine sediments underlying anoxic bottom water. Limnol Oceanogr. 1992;37(4):725–38.

Roden EE, Tuttle JH. Inorganic sulfur turnover in oligohaline estuarine sediments. Biogeochemistry. 1993 Jun;22(2):81–105.

Ruppert EE, Fox RS, Barnes RD. Invertebrate zoology: a functional evolutionary approach. 7th ed. Belmont, CA: Thomson-Brooks/Cole; 2004.

Sanford LP, Suttles SE, Halka JP. Reconsidering the physics of the Chesapeake Bay estuarine turbidity maximum. Estuaries. 2001 Oct 1;24(5):655–69.

Saunders JK, Fuchsman CA, McKay C, Rocap G. Complete arsenic-based respiratory cycle in the marine microbial communities of pelagic oxygen-deficient zones. Proc Natl Acad Sci. 2019 May 14;116(20):9925–30.

Schubel JR. Turbidity maximum of the northern Chesapeake Bay. Science. 1968;161(3845):1013–5.

Shibata A, Kogure K, Koike I, Ohwada K. Formation of submicron colloidal particles from marine bacteria by viral infection. Mar Ecol Prog Ser. 1997;155:303–7.

Shibl AA, Isaac A, Ochsenkühn MA, Cárdenas A, Fei C, Behringer G, et al. Diatom modulation of select bacteria through use of two unique secondary metabolites. Proc Natl Acad Sci. 2020 Nov 3;117(44):27445–55.

Shu Q. Molecular progresses of marine Planctomycetes: A review. Afr J Microbiol Res [Internet]. 2011 Dec 31 [cited 2023 Oct 26];5(33). Available from:

http://www.academicjournals.org/AJMR/abstracts/abstracts/abstract%202011/December%20Special%20Review/Shu%20et%20al.htm

Simon M, Alldredge A, Azam F. Bacterial carbon dynamics on marine snow. Mar Ecol Prog Ser. 1990;65:205–11.

Simon M, Grossart H, Schweitzer B, Ploug H. Microbial ecology of organic aggregates in aquatic ecosystems. Aquat Microb Ecol. 2002 Jun 26;28(2):175–211.

Stief P, Kamp A, Thamdrup B, Glud RN. Anaerobic Nitrogen Turnover by Sinking Diatom Aggregates at Varying Ambient Oxygen Levels. Front Microbiol [Internet]. 2016 Feb 5 [cited 2019 Jul 31];7. Available from: http://journal.frontiersin.org/Article/10.3389/fmicb.2016.00098/abstract

Stocker R. Marine Microbes See a Sea of Gradients. Science. 2012 Nov 1;338(6107):628–33.

Tee HS, Waite D, Lear G, Handley KM. Microbial river-to-sea continuum: gradients in benthic and planktonic diversity, osmoregulation and nutrient cycling. Microbiome. 2021 Sep 20;9(1):190.

Testa JM, Kemp WM, Boynton WR. Season-specific trends and linkages of nitrogen and oxygen cycles in Chesapeake Bay: Linked oxygen and nitrogen trends. Limnol Oceanogr [Internet]. 2018 May 31 [cited 2018 Oct 18]; Available from: http://doi.wiley.com/10.1002/lno.10823

Tourlousse DM, Yoshiike S, Ohashi A, Matsukura S, Noda N, Sekiguchi Y. Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon sequencing. Nucleic Acids Res. 2017 Feb 28;45(4):e23.

Trojan D, Schreiber L, Bjerg JT, Bøggild A, Yang T, Kjeldsen KU, et al. A taxonomic framework for cable bacteria and proposal of the candidate genera Electrothrix and Electronema. Syst Appl Microbiol. 2016 Jul 1;39(5):297–306.

Tsubokura A, Yoneda H, Mizuta H 1999. Paracoccus carotinifaciens sp. nov., a new aerobic Gram-negative astaxanthin-producing bacterium. Int J Syst Evol Microbiol. 1999;49(1):277–82.

Turk V, Malkin SY, Celussi M, Tinta T, Cram J, Malfatti F, et al. Ecological Role of Microbes: Current Knowledge and Future Prospects. In: Malone TC, Malej A, Faganeli J, editors. Coastal Ecosystems in Transition: A Comparative Analysis of the Northern Adriatic and Chesapeake Bay. John Wiley & Sons; 2021.

Turner JS, St-Laurent P, Friedrichs MAM, Friedrichs CT. Effects of reduced shoreline erosion on Chesapeake Bay water clarity. Sci Total Environ. 2021 May 15;769:145157.

Vanwonterghem I, Evans PN, Parks DH, Jensen PD, Woodcroft BJ, Hugenholtz P, et al. Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. Nat Microbiol. 2016 Dec;1(12):16170.

1290 Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, et al. Environmental
1291 Genome Shotgun Sequencing of the Sargasso Sea. Science. 2004 Apr 2;304(5667):66–
1292 74.

1293 Větrovský T, Baldrian P. The Variability of the 16S rRNA Gene in Bacterial Genomes and Its
1294 Consequences for Bacterial Community Analyses. PLoS ONE. 2013 Feb
1295 27;8(2):e57923.

1296 Wang H, Zhang C, Chen F, Kan J. Spatial and temporal variations of bacterioplankton in the
1297 Chesapeake Bay: A re-examination with high-throughput sequencing analysis. Limnol
1298 Oceanogr. 2020;65(12):3032–45.

1299 Wang J, Hood RR. Modeling the Origin of the Particulate Organic Matter Flux to the Hypoxic
1300 Zone of Chesapeake Bay in Early Summer. Estuaries Coasts [Internet]. 2020 Jul 25
1301 [cited 2021 Jan 27]; Available from: http://link.springer.com/10.1007/s12237-020-00806-
1302 0

1303 Weber T, Bianchi D. Efficient Particle Transfer to Depth in Oxygen Minimum Zones of the
1304 Pacific and Indian Oceans. Front Earth Sci [Internet]. 2020 [cited 2020 Dec 23];8.
1305 Available from: http://www.frontiersin.org/articles/10.3389/feart.2020.00376/full

1306 Willis A, Martin BD, Trinh P, Barger K, Bunge J. breakaway: Species Richness Estimation and
1307 Modeling [Internet]. 2018. Available from: https://adw96.github.io/breakaway/

1308 Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-
1309 species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Res. 2014 Jan
1310 1;42(D1):D643–8.

1311 Zhang Q, Blomquist JD. Watershed export of fine sediment, organic carbon, and chlorophyll-a
1312 to Chesapeake Bay: Spatial and temporal patterns in 1984–2016. Sci Total Environ.
1313 2018 Apr 1;619–620:1066–78.

1314 Zhang X, Roman M, Kimmel D, McGilliard C, Boicourt W. Spatial variability in plankton biomass
1315 and hydrographic variables along an axial transect in Chesapeake Bay. J Geophys Res.
1316 2006;111(C5):C05S11.

1317 Zhang Y, Xiao W, Jiao N. Linking biochemical properties of particles to particle-attached and
1318 free-living bacterial community structure along the particle density gradient from
1319 freshwater to open ocean: Linking POM to Particle-Attached and Free-Living Bacteria. J
1320 Geophys Res Biogeosciences. 2016 Aug;121(8):2261–74.

1321 Zhao R, Hannisdal B, Mogollon JM, Jørgensen SL. Nitrifier abundance and diversity peak at
1322 deep redox transition zones. Sci Rep. 2019 Jun 14;9(1):8633.

1323 Zimmerman AR, Canuel EA. Bulk Organic Matter and Lipid Biomarker Composition of
1324 Chesapeake Bay Surficial Sediments as Indicators of Environmental Processes. Estuar
1325 Coast Shelf Sci. 2001 Sep;53(3):319–41.

1326 Zumft WG. Cell Biology and Molecular Basis of Denitrification. MICROBIOL MOL BIOL REV.
1327 1997;61.

Supplemental Methods

Particle Processing

Phase 1. Separation of particulate material into size classes:

Water was gravity filtered, in sequence, through nylon mesh of decreasing pore size. The water that passed through the mesh was retained. Pore sizes were 500, 180, 53, 20, 5 µm and mesh diameters were ~142 mm. After filtration, each nylon mesh was back rinsed with ~500 ml of prefiltered "rinse water" to produce a resuspension of particulate matter from particles from each size class. The "rinse water" had been generated during transit by pumping surface water in sequence through water filters of size 10, 5, 1 µm to remove particles, followed by a 0.2 µm filter (Pall AcroPak 1500 Capsule with a Supor Polyethersulfone membrane) capsule which removes bacteria. After back-rinsing, the resuspended particles were split, with 45 mL saved for microscopy, ~200 mL used for particulate matter mass measurements and ~200 mL for DNA measurements. In all cases the actual volumes were carefully recorded and used for normalization during analysis. Water that had passed through all of the filters was also saved, 45 mL preserved for microscopy, ~1 L ml used for particulate matter mass measurements and ~1 L ml for DNA measurements.

Phase 2: Preservation of particulate material for microscopy, POM and DNA analysis

Microscopy samples were preserved with the addition of 1 mL of saturated, 0.2 µm filtered formaldehyde, and stored at -80°C.

To collect particulate matter for mass and elemental measurements, the resuspended particulate matter from each sample and size class was collected by vacuum filtration through a 1.2 µm nominal pore size, 25 mm diameter, GF/C glass fiber filter (Whatman WHA1822025). These filters had been previously pre-combusted at 400°C, and then pre-weighed using a Sartorius micro balance to a precision of 1 µg. To collect particulate matter in the 1.2 – 5 µm range, ~500 ml of the water that had gone through all of the nylon mesh was also collected on the 1.2 µm GF/C filter.

To collect DNA, resuspended particulate matter from each nylon filter was filtered through a 1.2 µm pore-size 47 mm diameter, polyethersulfone filter (Sterlitech PES1247100). DNA from the 1.2 – 5 µm and 0.2 – 1.2 µm ranges were collected by passing ~1 L of the water that had passed through the other filters, in series, through 1.2 µm and 0.2 µm pore size, 47 mm diameter, polyethersulfone filters using a peristaltic pump (1.2 µm filter as above, 0.2 µm filter — Sterilitech PES0247100). All DNA filters were folded, placed in a 2 mL sterile screw-cap cryogenic tube and flash frozen in liquid nitrogen. Samples were then stored at -80 °C prior to extraction.

Isotopic Analysis

At the University of California Davis Stable Isotope Facility, the following protocol was followed: Samples were combusted at 1080 °C in a reactor packed with chromium oxide and silvered copper oxide. Oxides were removed in a reduction reactor containing reduced copper at 650°C. The liberated gases were retained in a helium carrier, which flowed through a water trap containing magnesium perchlorate and phosphorous pentoxide. Carbon dioxide was temporarily stored in an adsorption trap until dinitrogen gas is analyzed, and then the carbon was analyzed.

Both carbon and nitrogen were quantified in an Elementar Vario EL Cube coupled with an Isoprime VisION isotope ratio mass spectrometer. The total carbon and nitrogen content of each filter were measured, as well as $\delta^{13}C$ and $\delta^{15}N$ relative to Vienna Pee Dee Belemnite and Air standards respectively.

Microscopy measurements of bacterial abundance

In cases where particle and microbial abundance was high enough that it would be challenging to enumerate cells directly on a microscope due to overlap between bacteria, fixed samples were first diluted 100-fold in 0.2 µm filtered deionized water, while in others sample water was used directly. Microbial abundances on each particle size fraction were measured by adding 1 µl/mL of Triton-X-100 per mL of sample, vortexing and then sonicating the sample for 30 seconds. Samples were then collected by vacuum filtration on a 0.2 µm pore size, 25 mm diameter, black PCTE track etched polycarbonate filter (Thomas Scientific # 4663H09). Samples were stained in some cases by incubating with a 20 mg/L DAPI solution for 5 minutes, in which case the DAPI was removed by filtration and the filters placed on a microscope slide, covered with non-fluorescent immersion oil and a coverslip. In other cases, samples were instead stained by placing Invitrogen ProLong Gold Antifade Mountant with DAPI solution (Fisher Scientific #P36941) under the coverslip, instead of immersion oil. The two methods gave similar results. Microbial abundances were counted on a Zeiss Axio Imager M2 epifluorescence microscope with a gridded eyepiece and microbial counts normalized to abundance (following Patel et al. 2007).

DNA Extraction

Frozen samples were combined with 400 µL lysis buffer (1% Sodium Dodecyl sulfate in filter sterilized STE (10mM Tris-Cl, 100mM NaCl, 1mM EDTA; pH 8.0)). Samples were heated at 95 °C for two minutes, put in a bead beater (BeadBug Microtube Homoginizer), without added beads, for 30 seconds. This process was followed by another round of heating, one more round of beating, and a final round of heating. We then added 400 µL of phenol-chloroform solution (25 parts phenol solution (pH8; Sigma P4557): 24 parts chloroform (Sigma C2432); 1 part isoamyl alcohol (Sigma W205702)). Samples then underwent one additional beating step (30 seconds). During this final beating step, the filters dissolved in the phenol solution. The sample was then centrifuged at 13,000 g at room temperature (~20 °C) for two minutes to separate the phenol layer from the aqueous layer. The samples were again placed in the bead beater for an additional 30 seconds and then re-centrifuged to increase the samples' exposure to phenol chloroform. The remaining phenol layer was viscous due to the presence of dissolved filter and was removed with a pipette from which the end had been cut off to widen the opening. An additional 400 µL phenol chloroform was added, samples were vortexed for 30 seconds to mix, the sample was re-centrifuged and the phenol again removed. Next 400 µL of chloroform-isoamyl alcohol (24:1 ratio) solution was added and the samples again vortexed and centrifuged again to separate the layers. At this point the aqueous layer was transferred to a new tube.

To precipitate DNA, to the aqueous layer 10 µg of glycogen (a co-preciptant) and 0.31 times the lysate volume of 10.5 M filter sterilized Ammonium Acetate was added. The sample was mixed by inverting 5 times, and then 1.31 times the lysate volume of isopropyl alcohol was added. Samples were incubated at -20°C overnight (~10 hours). DNA was pelleted at the

bottom of the tube by spinning at 15,000 g for 60 minutes. The supernatant was decanted off. Then the pellet was rinsed with 250 mL -20 °C ethanol and the pellet spun again for 30 minutes. The pellet was then air dried in an inverted tube (~2 hours). The pellet was resuspended in 25 µL of TE (10 mM TRIS-HCl, 1 mM EDTA, pH 8) and incubated at 37 °C for 1-2 hours for the DNA to dissolve. DNA concentration was measured using a qubit fluorometer. A sample aliquot was then diluted to a final concentration of 2 ng/ul. Samples were then preserved at -80 °C prior to further analysis.

Amplicon Libraries

Our amplicon libraries leveraged the 515FY-926R primers (Parada et al. 2016) which had been amended with overhangs for a final sequence of

Overhang + **515FY**:
5’ TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**GTGYCAGCMGCCGCGG** 3’

Overhang + **926R**:

5’ GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**CCGYCAATTYMTTTRAGTTT** 3’

Samples were amplified by combining, for a final volume of 25 µL, 1X Accustart Master Mix; 0.3 mM of each of the forward and reverse primers; 0.5ng/ul of DNA and 5 * 10^4 of spike-in sequence (Genbank Accession LC140931). Samples were amplified in a thermal cycler (BioRad T-100) as follows: Samples underwent initial denaturation at 95 °C for 120s. There were then 25 cycles of denaturation (95 °C * 45 s) annealing (50 °C * 45 s), and elongation (68 °C, 90 s). Following the amplification cycles, there was a final elongation step of (68 °C, 300 s). Samples were then held at 10 °C for less than one hour before being refrigerated at 4 °C. We checked to ensure that amplicons appeared in our samples and were the appropriate length by running samples on an agarose gel. Samples were then cleaned using AMPure XP beads, following the manufacturers’ instructions with a bead:sample ratio of 0.8.

Samples were amplified to append unique forward and reverse barcodes and illumina adapters:

i7 index primer

5’ CAAGCAGAAGACGGCATACGAGAT**XXXXXXXX**GTCTCGTGGGCTCGG

i5 index primer

5’ AATGATACGGCGACCACCGAGATCTACAC**XXXXXXXX**CTCTCTAT

Where bold **X**s stand in for Illumina i7 index sequences (N701-N729) and i5 index sequences (S502-S522).

Samples were again amplified in 25 µL volumes with 1X Accustart master mix, 0.4 mM of each primer, and 7 µL of amplified DNA (we do not dilute to a common concentration between rounds). Samples were amplified as in round one, but with only five amplification cycles, rather than 30.

Fragments were again assessed by agarose gel and cleaned with AMPure XP beads. Samples were eluted with TE following amplification. Samples were diluted to a common concentration of 5 ng/ul and sent to IMET’s Bioanalytical Services laboratory (BASLab) for

sequencing. We found that no samples from station CB5.1 Bottom amplified. Those samples were thus re-amplified as part of another batch of samples. Those samples followed the same protocol but were eluted in EBT (10 mM Tris, pH 8, 0.1% Tween20) for the final elution step. Those samples were sent to UC Davis's Core DNA lab for sequencing. Both sequencing facilities ran our samples on an Illumina MiSeq and returned demultiplexed samples to us.

Two mock community samples of the genomic DNA from 20 organisms, one in which each species had even proportions and one in which they had staggered proportions (BEI Resources HM-782-D and HM-783D), were run alongside the environmental samples for each run. We also ran a "Generous Donor" sample, DNA which we collected from the Horn Point Laboratory Pier and which we include in all runs so that they can be cross compared.

Amplicon Bioinformatics

To process amplicon sequence data, primers were removed from sequences with Cutadapt (Martin 2011). Due to relatively low sequence quality (see Supplemental Results; Amplicon Processing), we used the DADA2 `filterandtrim()` algorithm to retain sequences with fewer than three errors on the forward read and fewer than five errors on the reverse read, and truncated sequences to 230 sequences in the forward and 220 in the reverse direction. DADA2 was then used to learn error rates, dereplicate identical sequences, call amplicon sequence variants, and to merge forward and reverse reads. Samples from the BasLAB and DNATECH sequence libraries were called separately, as recommended by the package developers, and were then merged after sequence calling. Chimeric sequences were then removed. Taxonomy was called using DADA2's `assignTaxonomy()` function using the Silva database, version 132 (Quast et al. 2013; Yilmaz et al. 2014), which had been modified by adding the spike in sequences. This scheme allowed us to classify bacterial, archaeal and eukaryotic sequences, as well as spike sequences.

Supplemental Results

Hydrographic Context

Hydrological conditions can impact the delivery of particles from rivers. The majority (90%) of terrestrial suspended solids that enter Chesapeake Bay come from the James, Potomac, and Susquehanna Rivers (Figure S1 for locations) (Zhang and Blomquist 2018). The sampled sites are primarily influenced by the Susquehanna River because the water flows from North to South and all of the sampled sites are north of the James River and only site CB5.5 is south of the Potomac River. Discharge at the Conowingo Dam plays a major role in determining of suspended sediments load of the water that enters the Bay from the Susquehanna River (Zhang et al. 2016) Thus, despite differences in hydrological conditions over the sampling period (i.e., CB5.1 and CB5.5 were sampled at baseflow conditions, CB4.3C and CB3.3C were sampled during the rising limb of a storm, and CB3.2 and CB3.1 were sampled on the falling limb of a storm; Figure S2A-C), discharge patterns at the Conowingo Dam did not substantially vary over the period of sampling (Figure S2A-C) and therefore it is likely that delivery of suspended sediments also did not greatly vary over the period of sampling. Additional information about physics, chemistry, and the particle size to mass and size to abundance distribution spectra of each site, when they were sampled for this study, are described in Dougherty et al. (2021).

Amplicon Processing

Mock community sequencing suggested that overall community structure was broadly in line with input communities. In the first sequencing run the even mock community, which contained 20 species each of which can have more than one 16S sequence, had 25 ASVs that comprised at least 1% of the community structure each for a total of 76.8% of the total mock community. These reads varied in relative abundance from 9.3% to 1.1%. There were 326 rare ASVs each comprising < 1% of the community. In the second run, there were 22 ASVs that comprised at least 1% of the community structure ranging in relative abundance from 8.9% to 1.3% each for a total of 96% of the overall community, and 189 rare ASVs.

Per sample sequence depth ranged from 7034 reads through 304399 reads (median = 91838 ± 38019 (median adjusted deviation, MAD)). We use an internal reference standard sample, a generous donor DNA sample collected at the HPL pier, as part of every sequencing run that our lab does. Sequence quality of this generous donor sample was characteristic of the other samples and allowed us to compare sequence quality between runs (Figure S7A-D) while ensuring that the community structure was similar between runs. The sequence quality of samples from BasLab appeared to drop off more quickly than other projects that we had encountered, with quality scores dropping below 30 after about 100 cycles for both the forward and reverse reads. In contrast, the samples run by DNATECH appeared to remain above 30 through 220 bases in the forward direction and 200 cycles in the reverse direction (Figure S7A-D).

DADA2's chimera checking program suggested that a high fraction of our reads were chimeras (range 7.2% - 53.4%; median 33.6% ± 9.4% MAD), in both of our runs. While high fractions of detected chimeras are often related to skipping primer trimming steps, we confirmed that this step did indeed occur for this analysis. Following all processing steps, we were left with between 2135 and 142572 reads (median 33803 ± 20198 MAD), including spike in sequences. After removing spike in sequences, all samples but three had more at least 9900 reads. The low read number samples were CB4.3 Surface 0.2 µm (852 reads); CB3.1 Bottom 500 µm (860 reads) CB3.3C Bottom 180 µm (1248 reads).

Microbial community structure data (Table S1), microbial taxonomic information (Table S2), and corresponding environmental parameters (Table S3) are available as supplemental tables.

Supplemental Figures

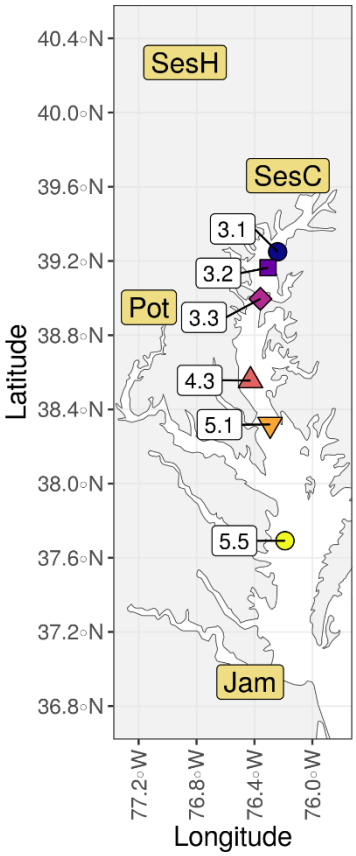


Figure S1. Map of the study area, reflecting not only the station locations shown in Figure 1A, but also the locations hydrological stations describing discharge by key rivers into the Chesapeake Bay (Figure S2). Each station is labeled in light goldenrod colored rectangles (SesH, Sesquihanna River at Harrisburg; SesC, Susquehanna River at Cantonsville; Pot, Potomic River near Washington DC). The James river (Jam), mentioned in the main text, is also shown.

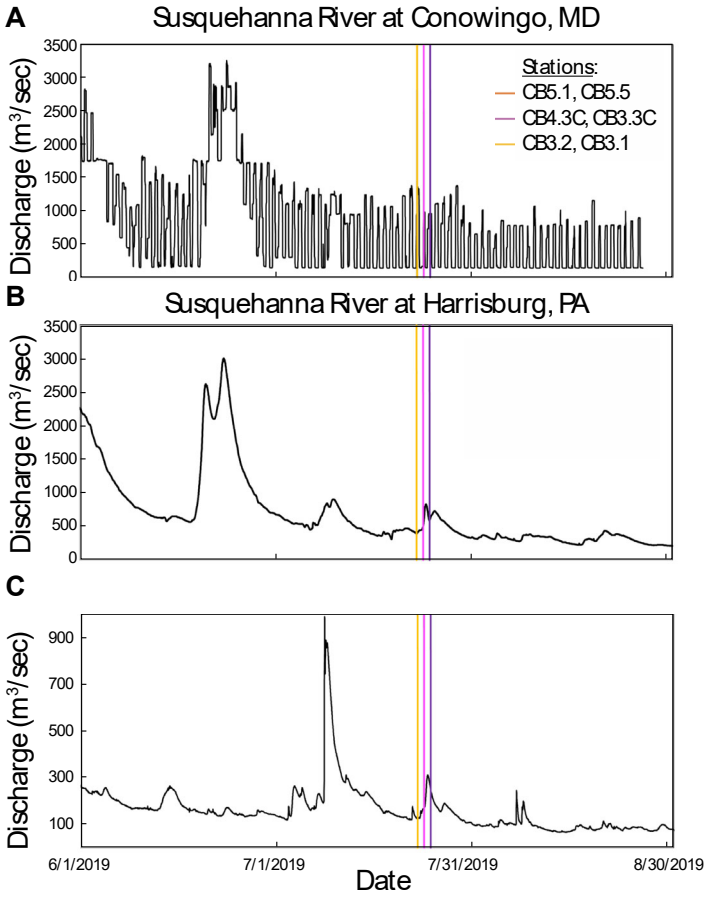
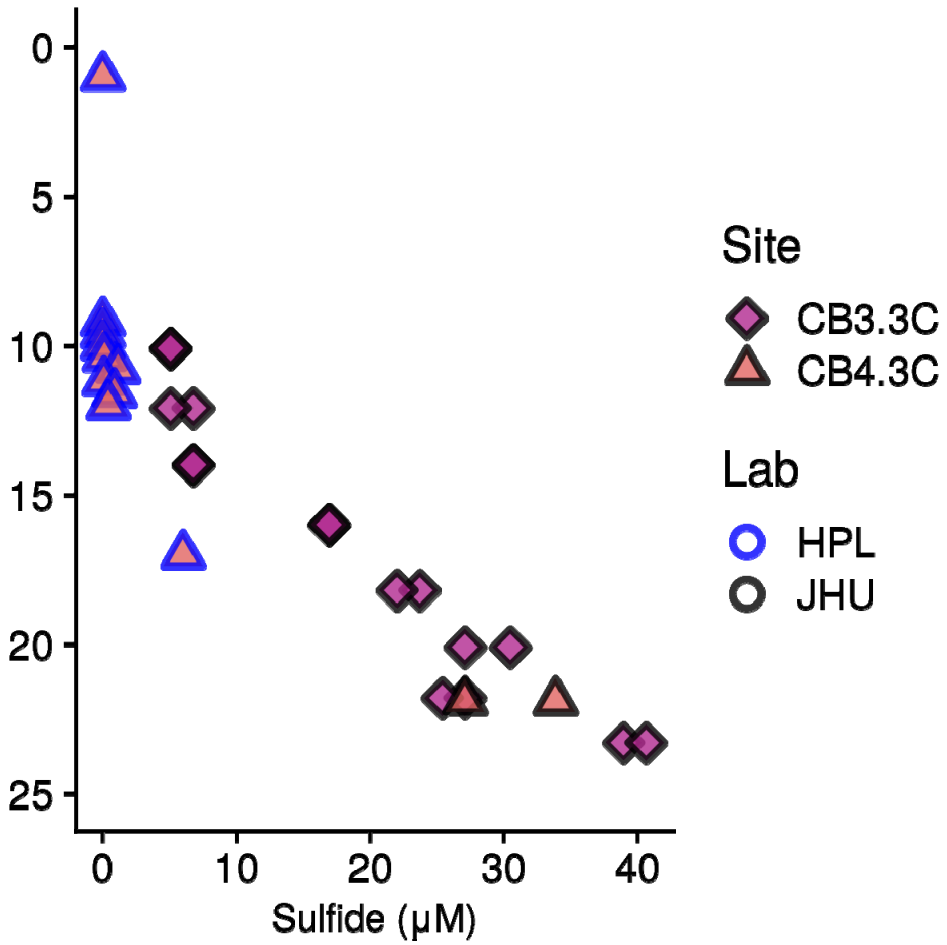


Figure S2. Discharge at United States Geological Survey (USGS) gauges at A) the Susquehanna River at Conowingo, MD (USGS 01578310), B) the Susquehanna River at Harrisburg, PA (USGS 01570500), B) the Susquehanna River at Conowingo, MD (USGS 01578310), and C) the Potomac River near Washington, DC (USGS 01646500). Data downloaded from https://waterdata.usgs.gov/.



1541

1542 Figure S3. Hydrogen sulfide concentrations, measured at stations CB3.3C and CB4.3C. The
1543 laboratory which performed the analysis differed between samples and are shown: Horn Point
1544 Laboratory (HPL) and Johns Hopkins University (JHU). Sulfide was not measured at any of the
1545 other stations, but sulfide was also not detected by smell at any of the other stations.

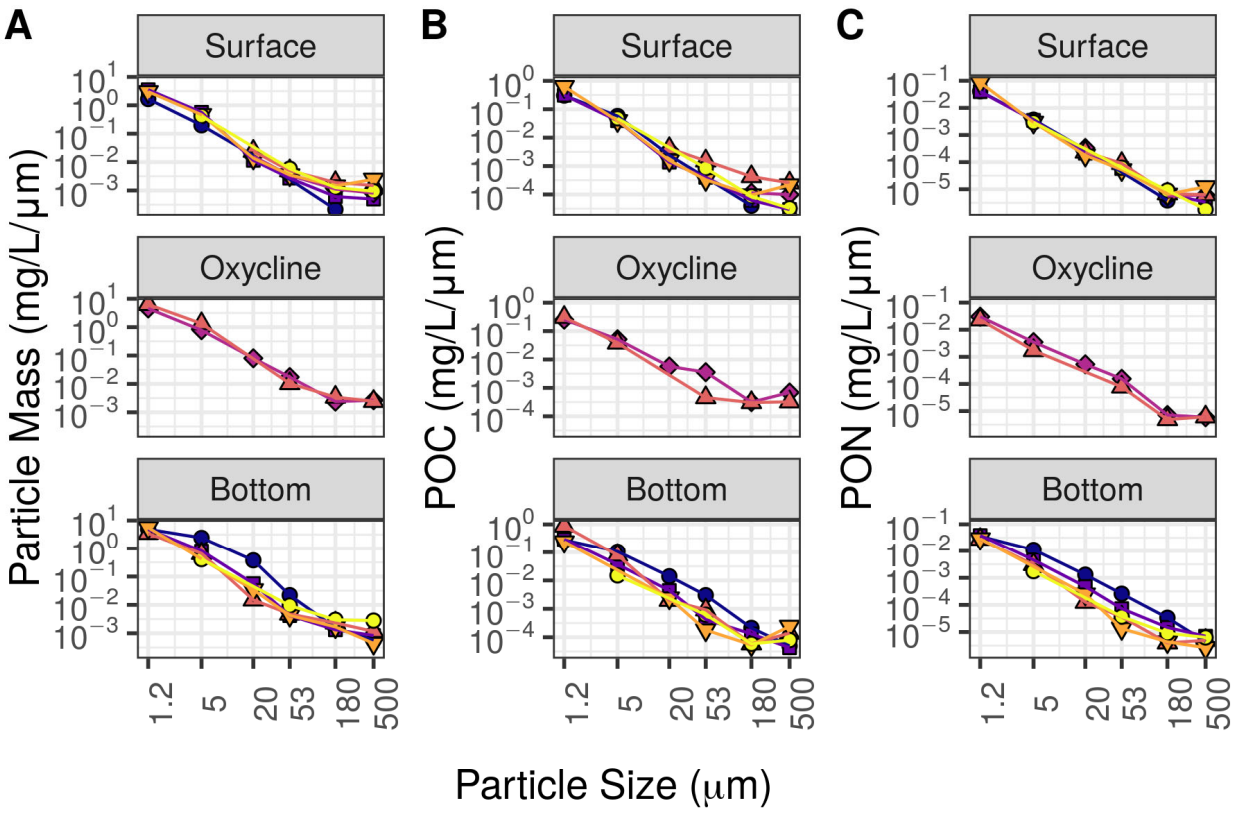


Figure S4. Particle (A) mass, (B) particulate organic carbon (POC) content and (C) particulate organic nitrogen (PON) content associated with each size fraction, at each measured depth, of each station. Colors and shapes of points correspond to stations as shown in Figure 1A and Figure S1.

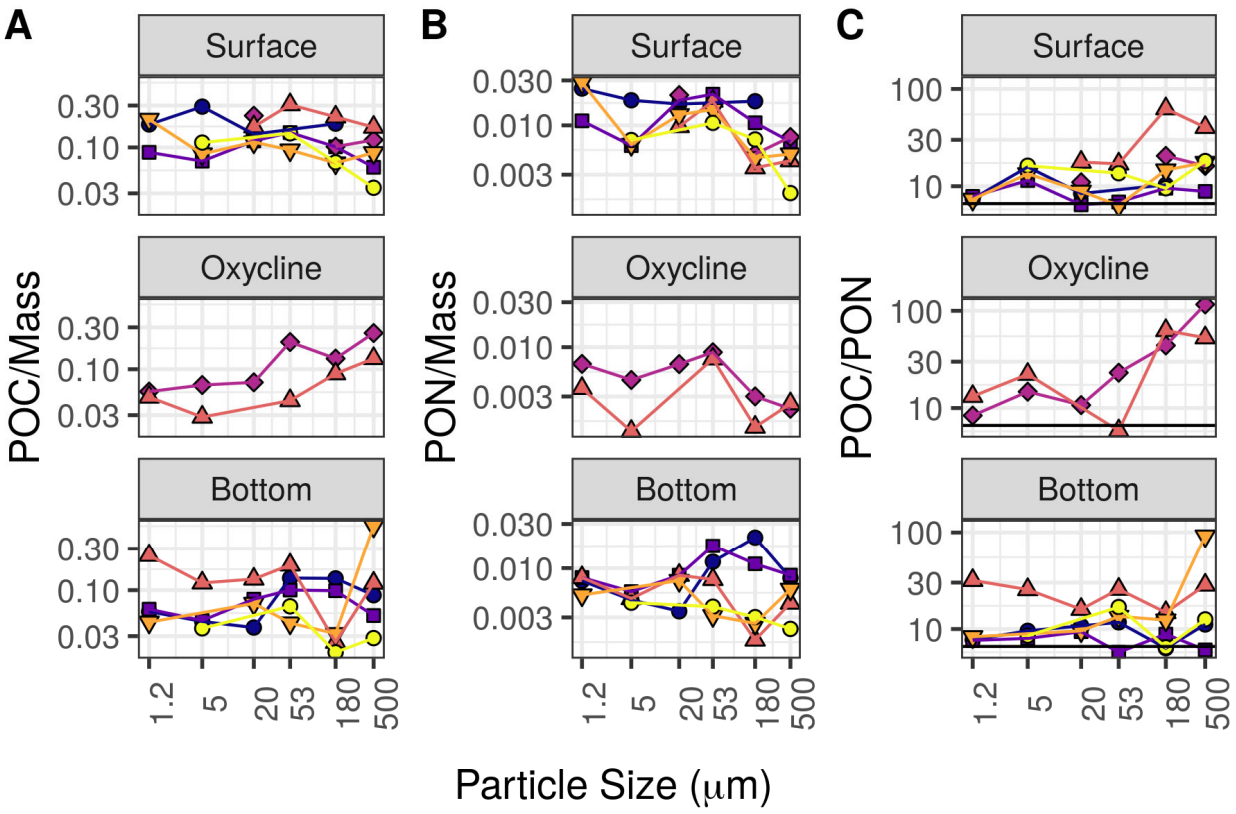


Figure S5. Ratios between the particle characteristics shown in Figure S2, numerators and denominators are all in the same units, and so y axis values are unitless. (A) Particulate organic carbon to mass ratios; (Particulate organic nitrogen to mass ratios); (C) Particulate organic carbon to nitrogen ratios. Colors and shapes of points correspond to stations as shown in Figure 1A.

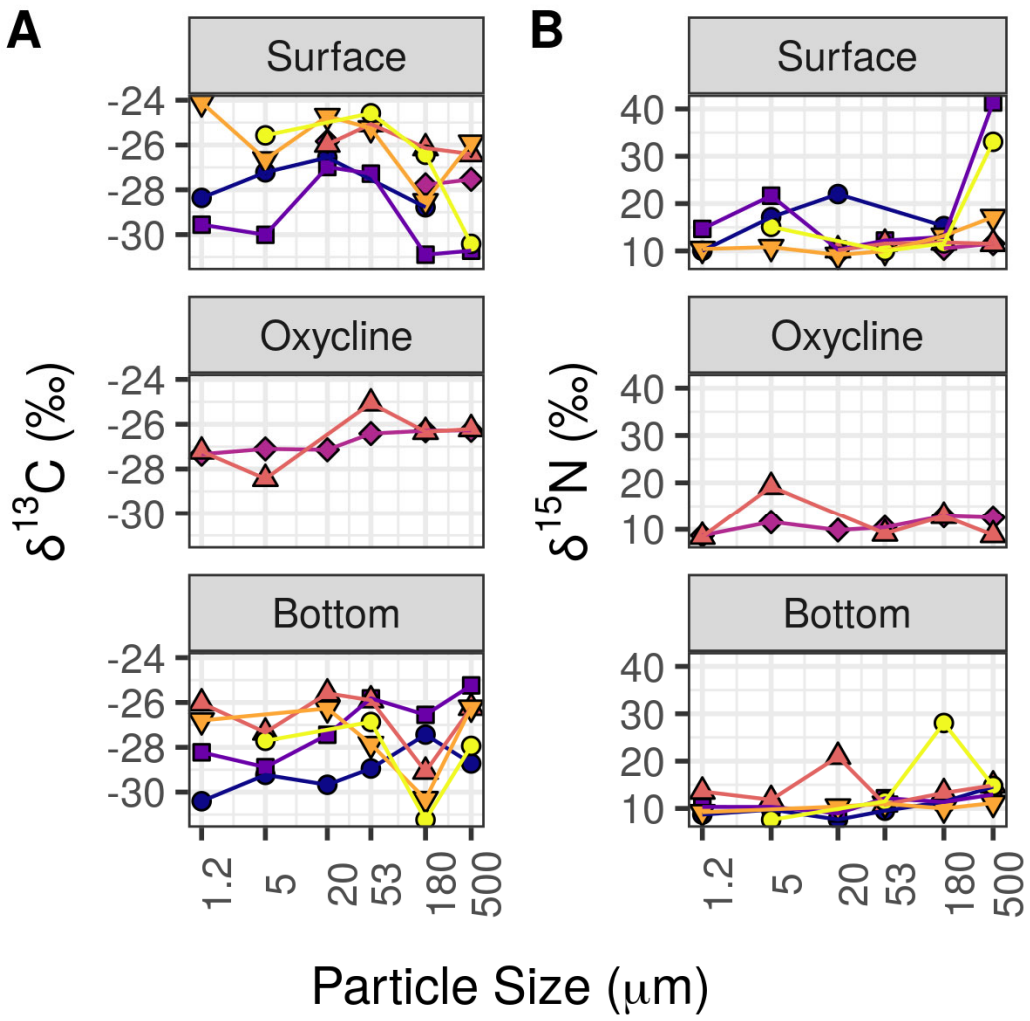


Figure S6. Particulate organic carbon and nitrogen isotopic ratios (**A**) $\delta^{13}C$ and (**B**) $\delta^{15}N$. Colors and shapes of points correspond to stations as shown in Figure 1A.

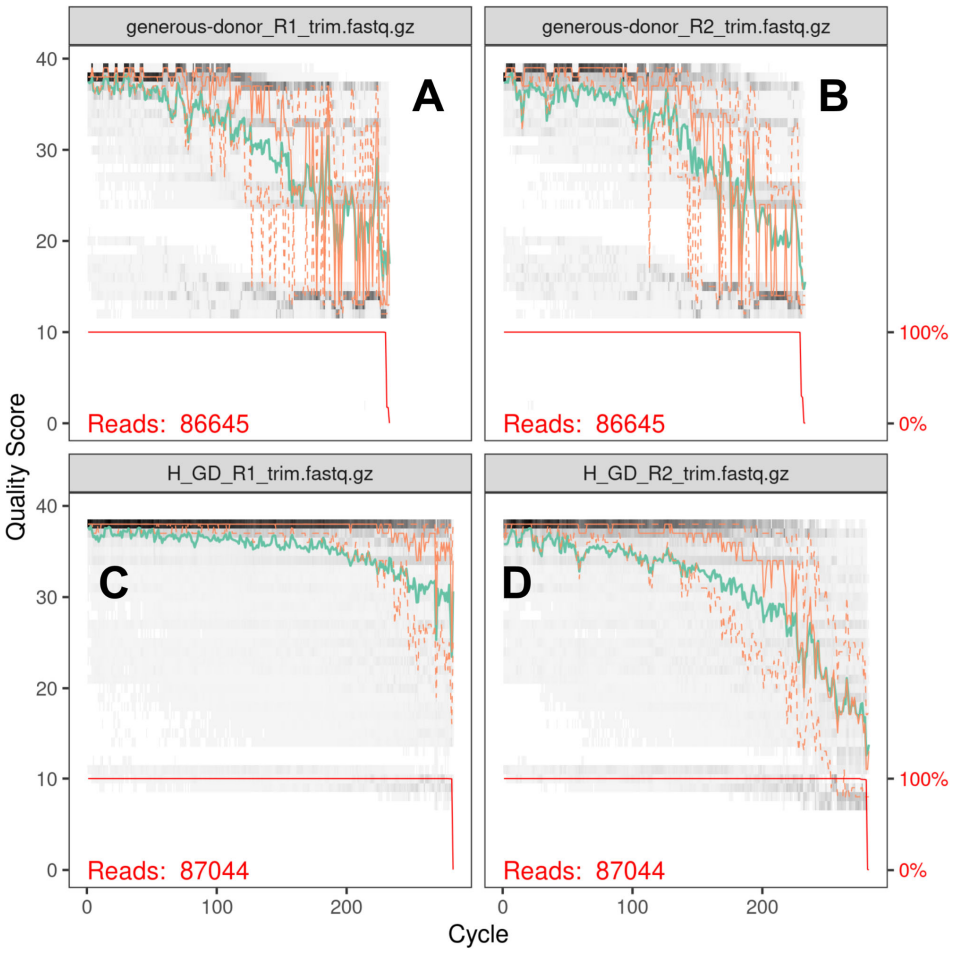


Figure S7. Read quality of generous donor samples (collected at the HPL pier, all size fractions) from our two different sequencing runs. **A-B.** most samples, sequenced by IMET BasLAB, **C-D.** samples from station 5.1 Bottom. Sequenced by UC Davis. **A,C.** Forward read quality. **B,D.** Reverse read quality. This is a standard output by the DADA2 program (Callahan et al. 2016). The heatmap indicates the distribution of scores, as well as a mean best fit score (line). The red line indicates the fraction of samples that are at least a given length.

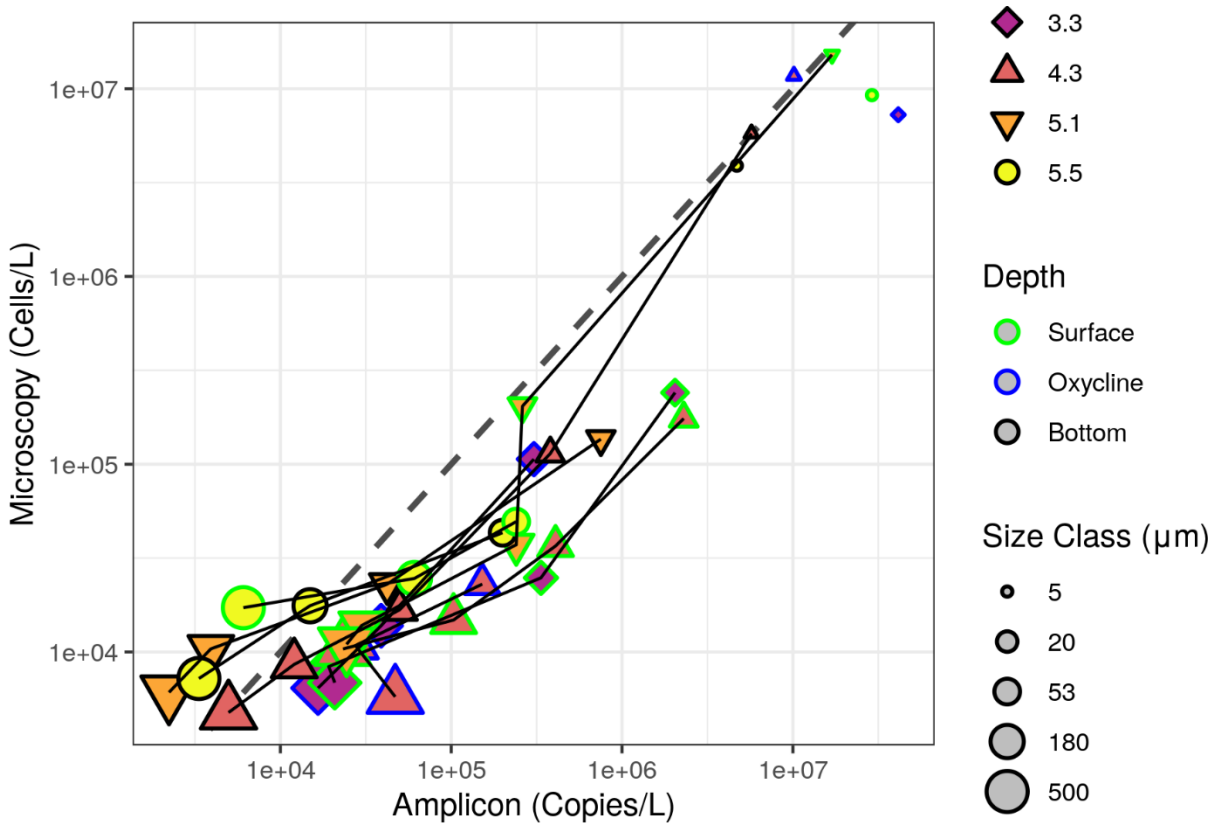


Figure S8. Microscopy-based estimates of microbial abundance compared to amplicon-based estimates. The black diagonal is the one to one line.

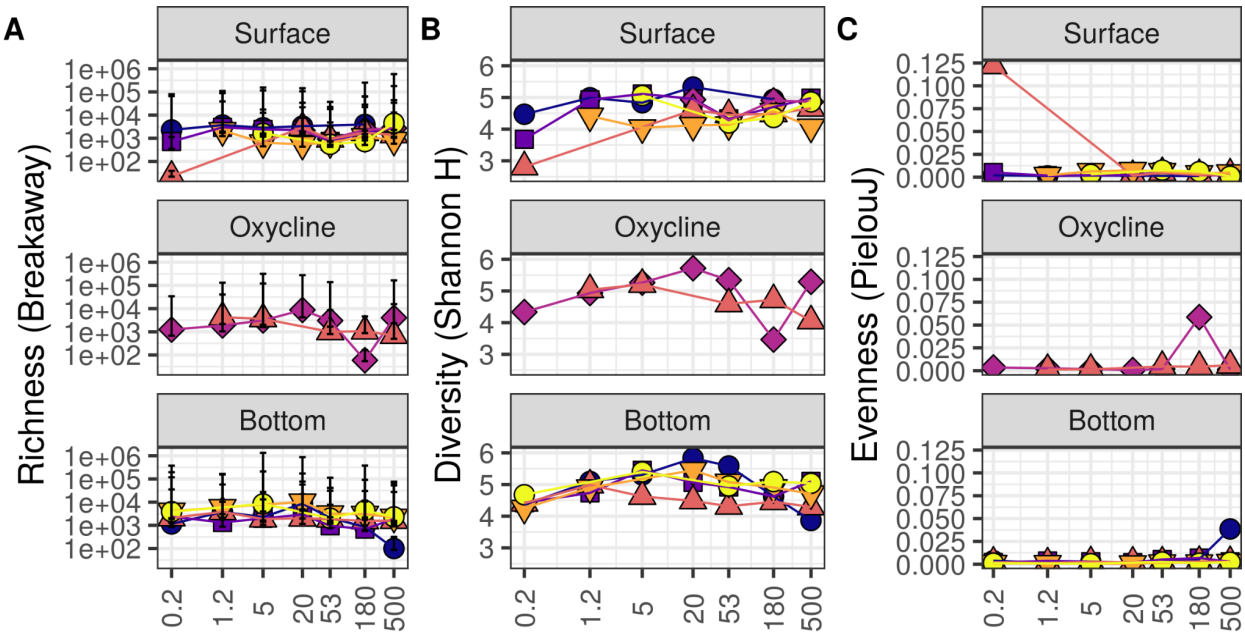


Figure S9. Estimates of **A** richness, **B** Shannon diversity and **C** evenness at each sample. Confidence intervals in *A* correspond to upper and lower bounds specified by the breakaway package. Colors and shapes of points correspond to stations as shown in Figure 1A. Note that richness, but not the other metrics, is on a log scale.

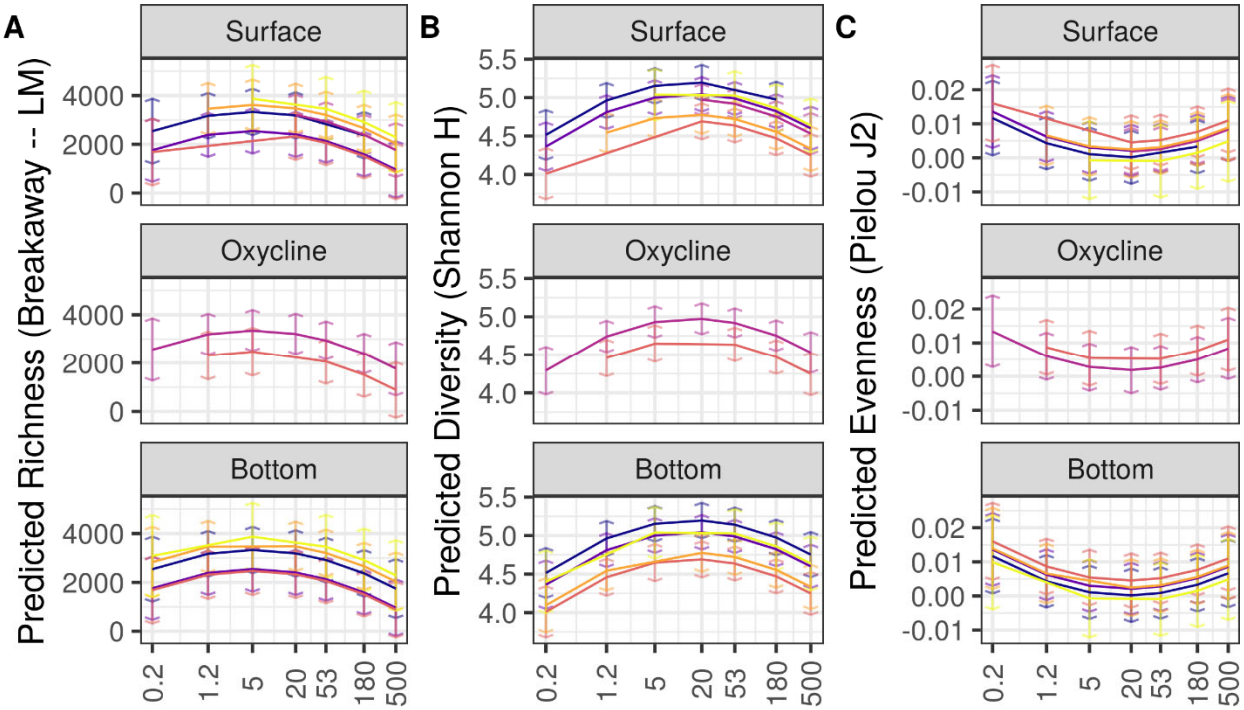


Figure S10. Linear model predicted outputs of **A** richness, **B** Shannon diversity and **C** evenness which show general trends with size and latitude at the surface and bottom of the

water column. Colors of the curves correspond to the colors shown in Figures 1A and S9. Confidence intervals are two standard errors of predicted mean values.

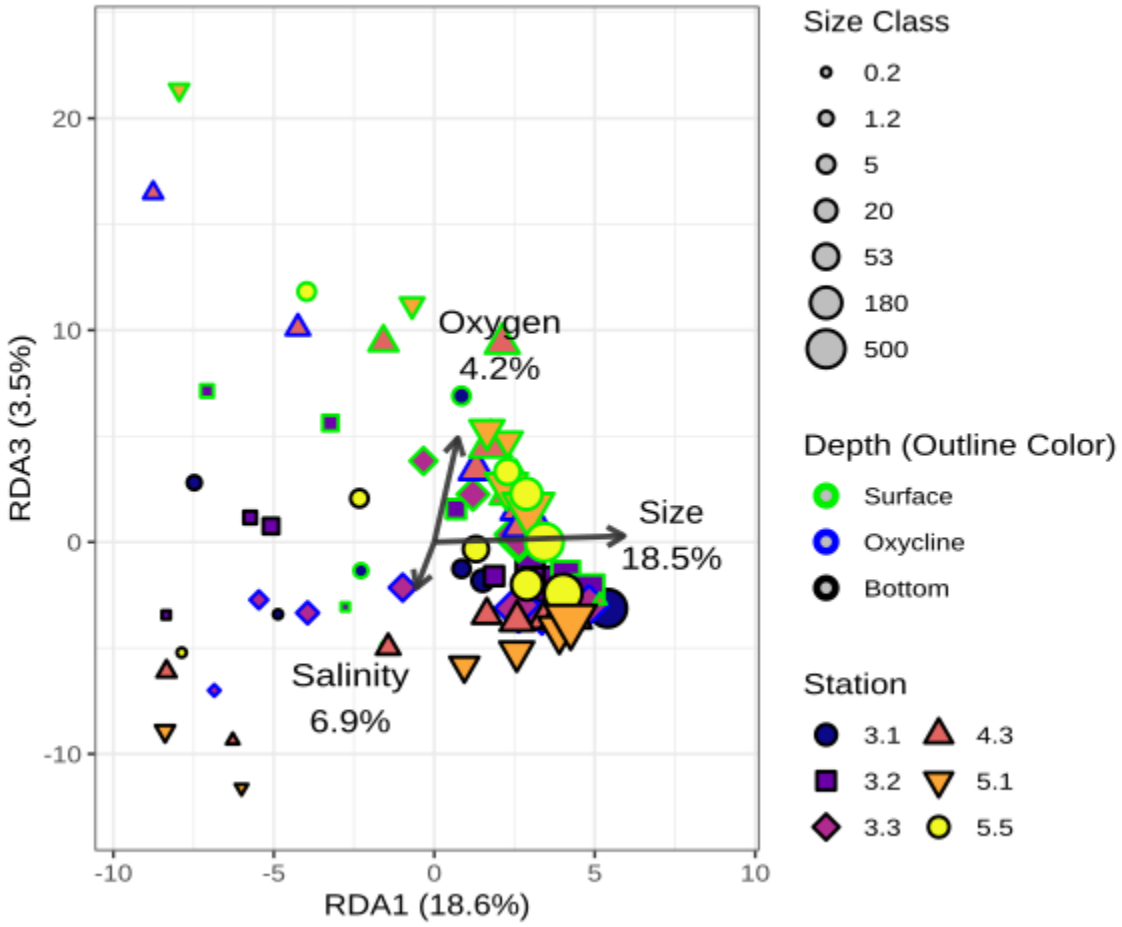


Figure S11. Axes RDA1 and RDA3 of the redundancy analysis shown in Figure 3.

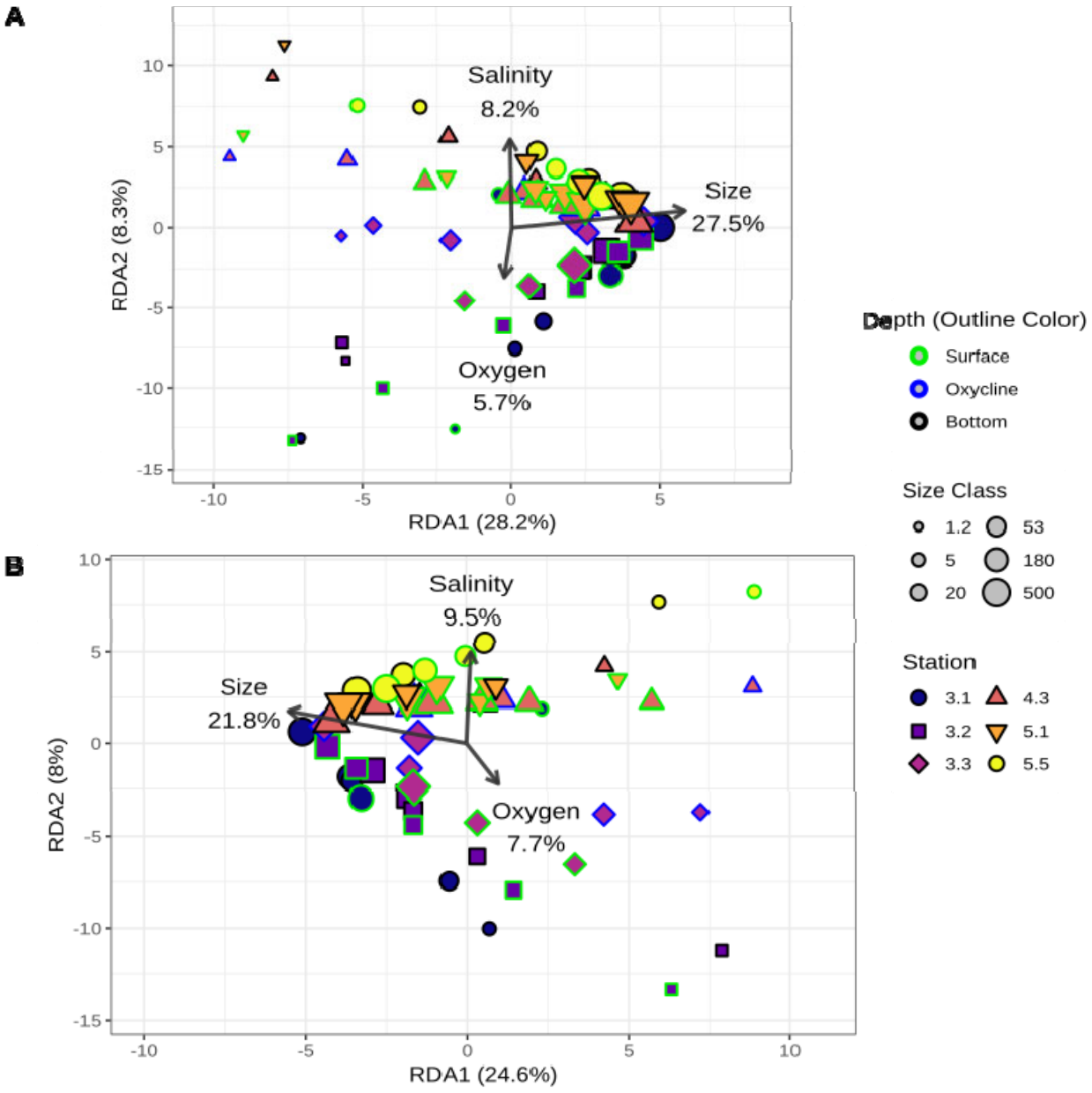
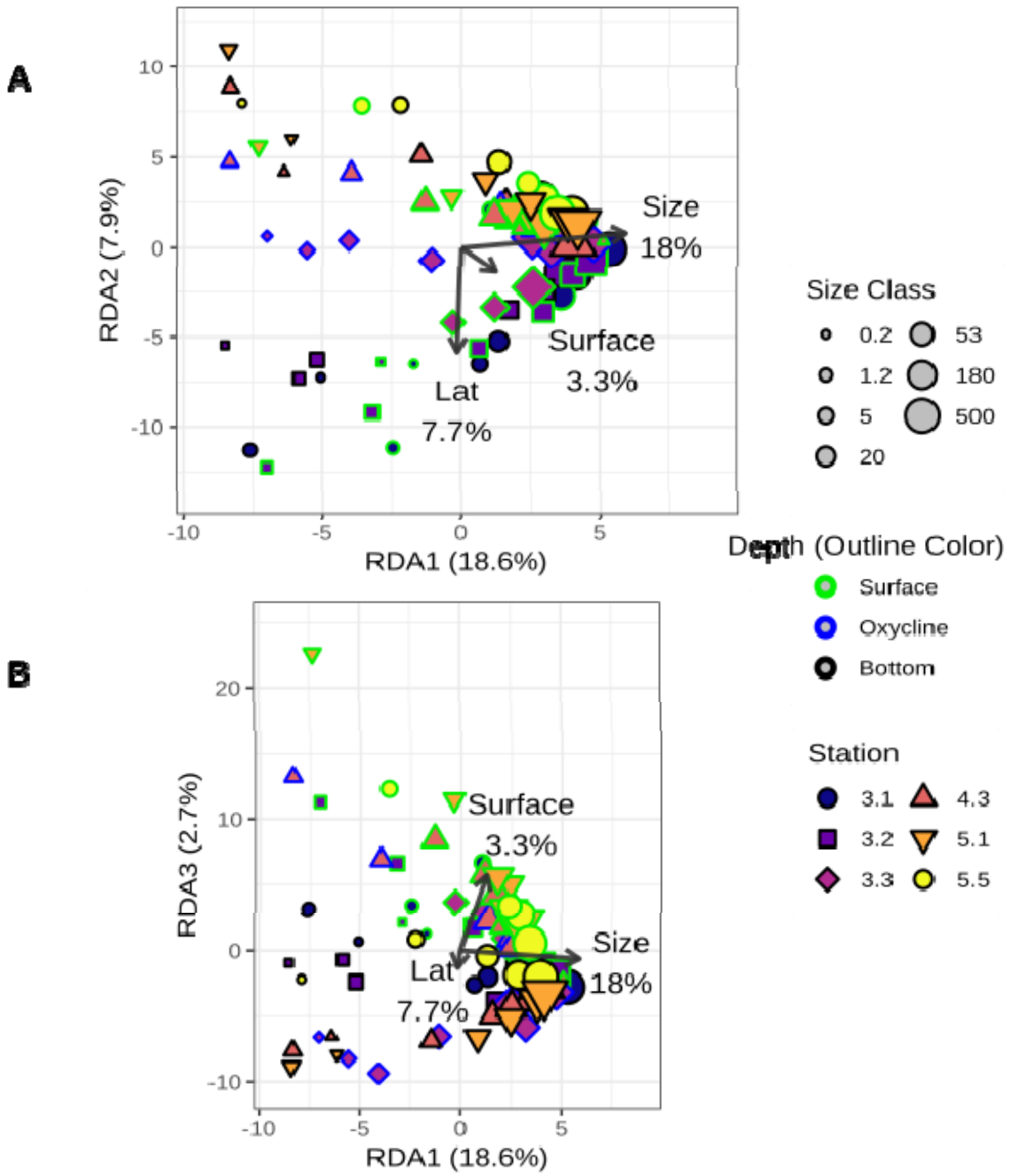
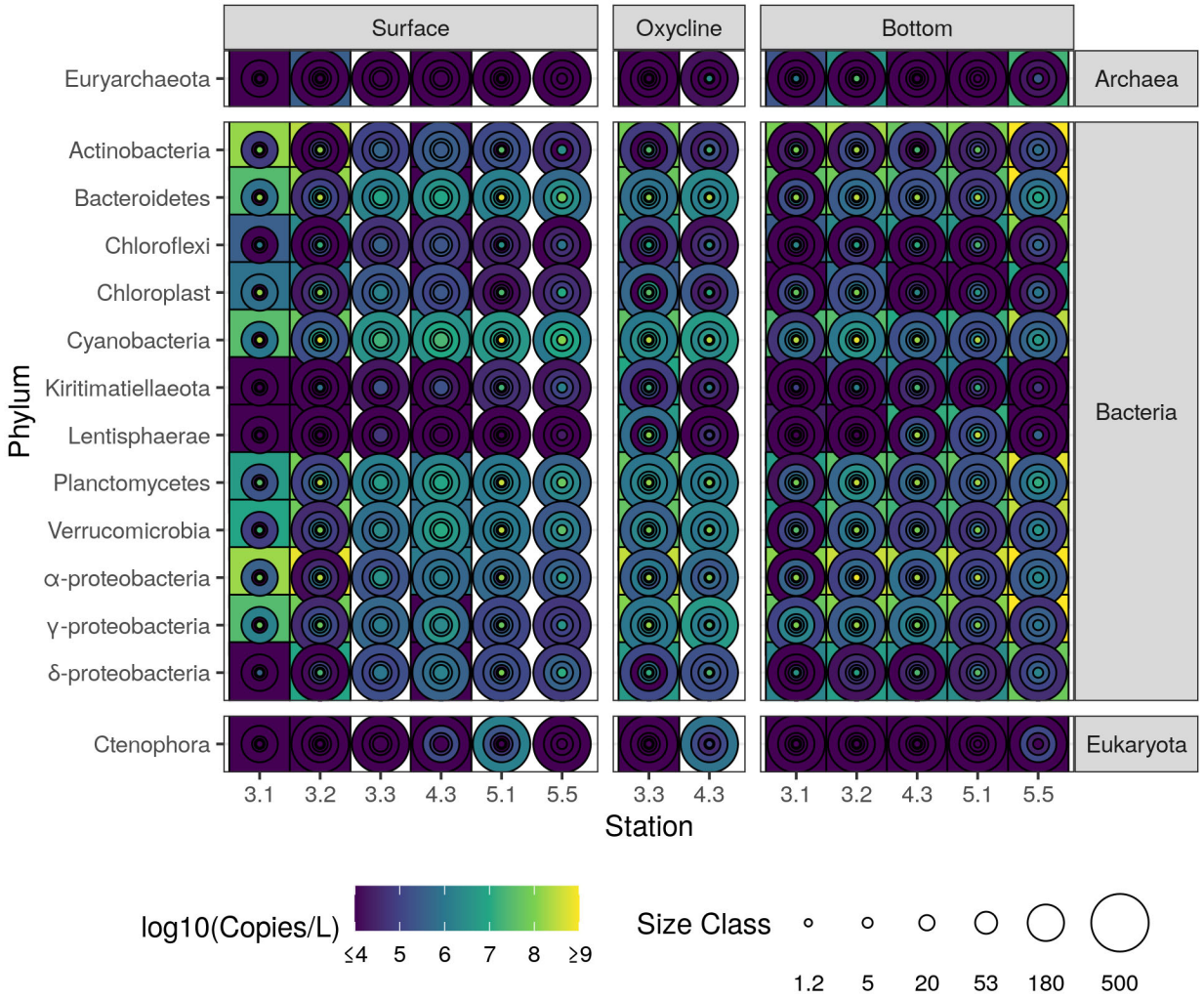


Figure S12. Redundancy analysis of the relationship between Salinity, Oxygen concentration and particle size microbial community structure (**A**) excluding samples from the (0.2 – 1.2 µm size classes and (**B**) excluding samples from the 0.2 – 1.2 µm and 1.2 – 5 µm size classes.



1605

1606 Figure S13. RDA in which latitude and whether the particle was collected at the surface, rather
1607 than salinity and oxygen, are used to predict community structure, as in Figures 3 and S11. A.
1608 RDA axes 1 and 2. B. RDA axes 1 and 3.

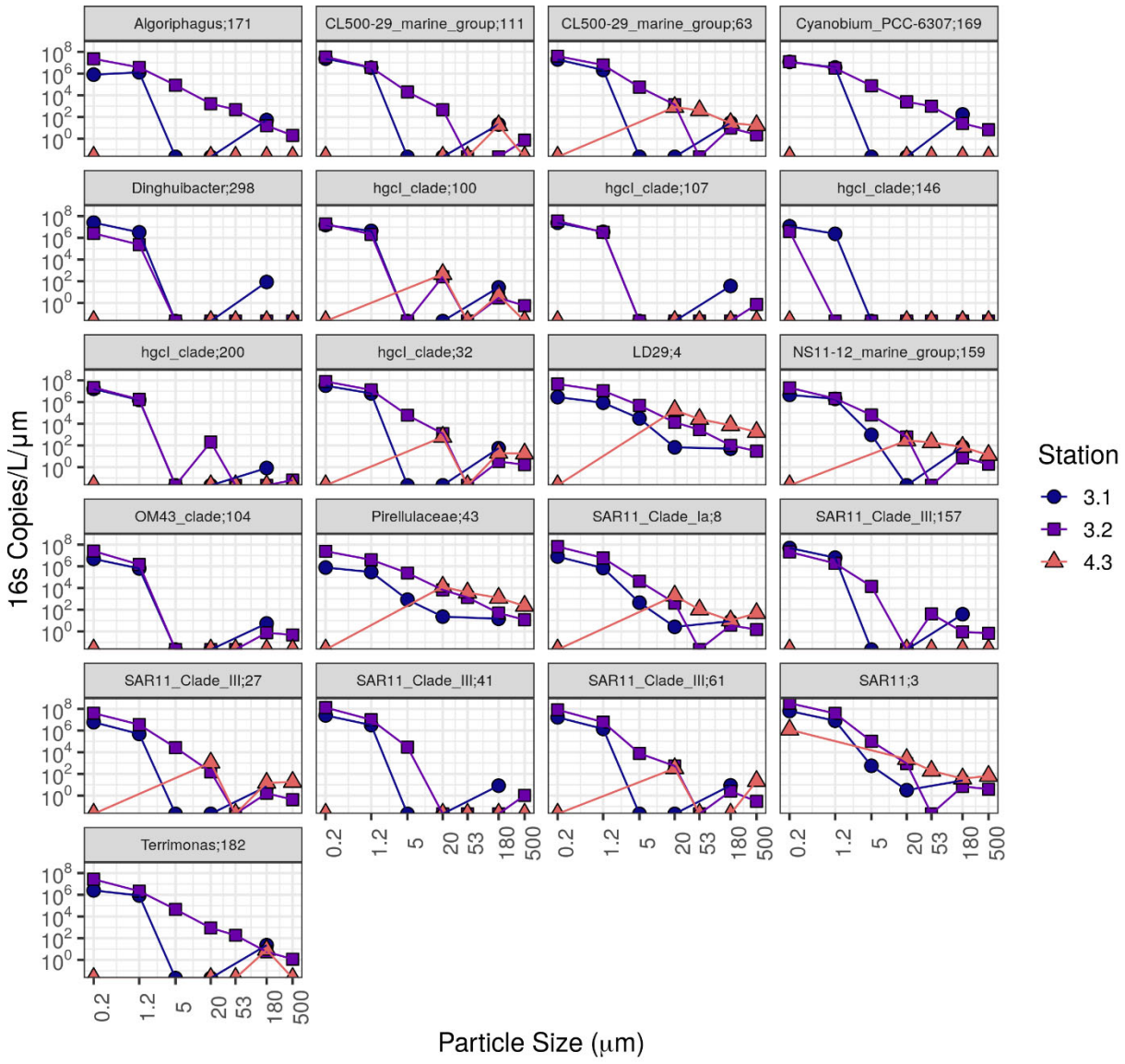


1609

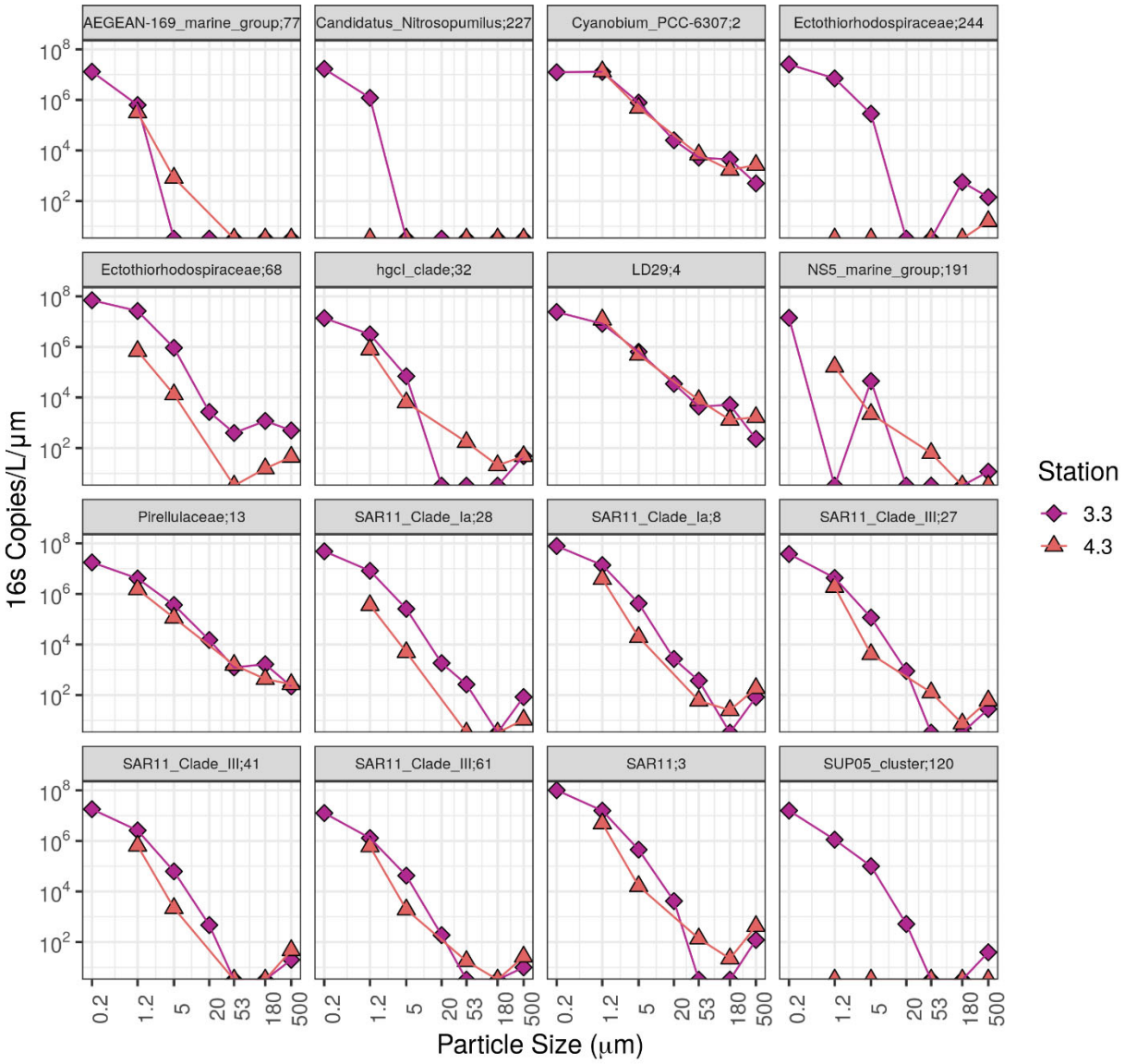
1610

1611 Figure S14. Phylum level abundance of bacteria, normalized to water volume, rather than
1612 particle mass.
1613

A

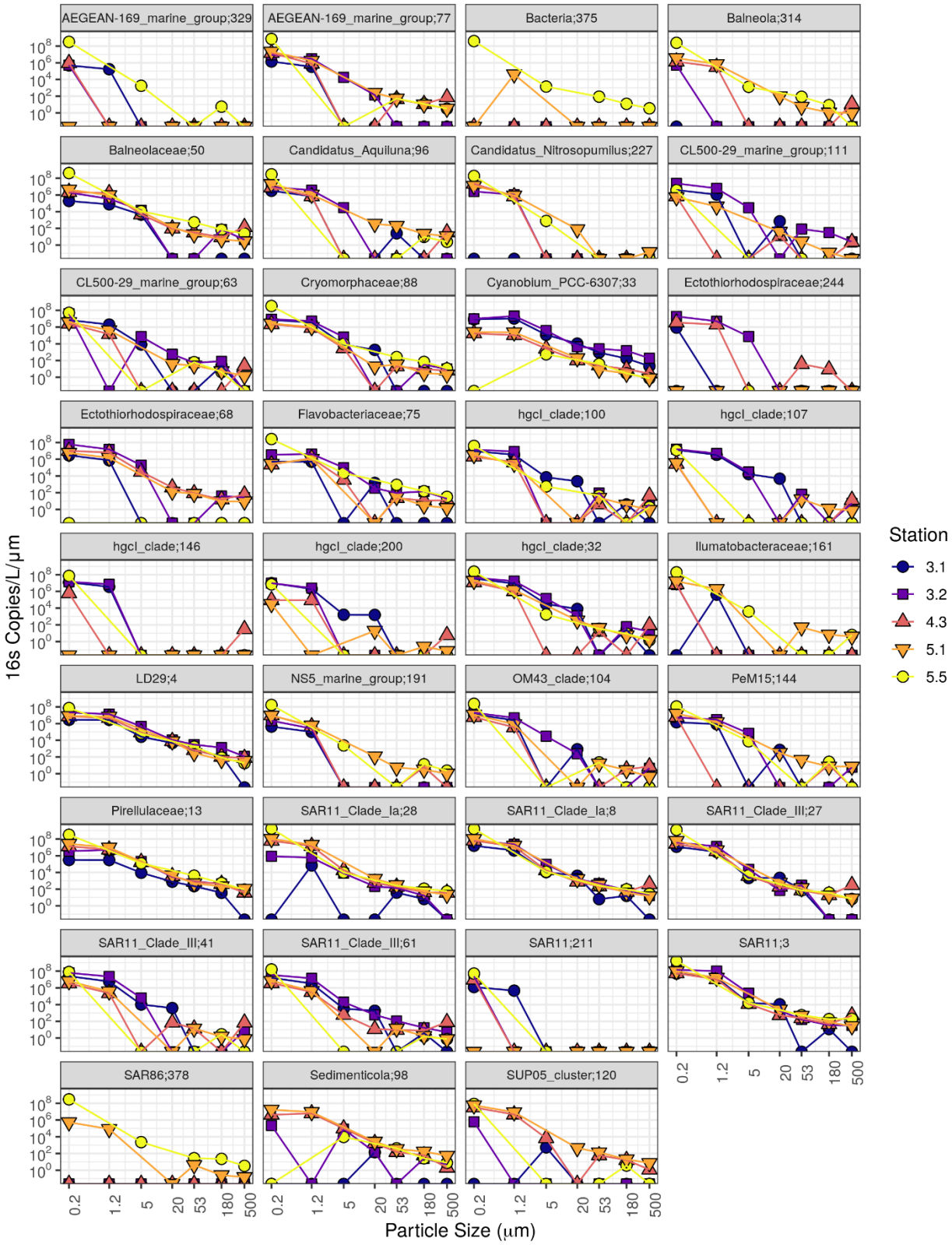

Algoriphagus;171
CL500-29_marine_group;111
CL500-29_marine_group;63
Cyanobium_PCC-6307;169
Dinghuibacter;298
hgcI_clade;100
hgcI_clade;107
hgcI_clade;146
hgcI_clade;200
hgcI_clade;32
LD29;4
NS11-12_marine_group;159
OM43_clade;104
Pirellulaceae;43
SAR11_Clade_Ia;8
SAR11_Clade_III;157
SAR11_Clade_III;27
SAR11_Clade_III;41
SAR11_Clade_III;61
SAR11;3
Terrimonas;182
16s Copies/L/µm
Particle Size (µm)
Station
3.1
3.2
4.3

1618 **B**


AEGEAN-169_marine_group;77
Candidatus_Nitrosopumilus;227
Cyanobium_PCC-6307;2
Ectothiorhodospiraceae;244
Ectothiorhodospiraceae;68
hgcI_clade;32
LD29;4
NS5_marine_group;191
Pirellulaceae;13
SAR11_Clade_Ia;28
SAR11_Clade_Ia;8
SAR11_Clade_III;27
SAR11_Clade_III;41
SAR11_Clade_III;61
SAR11;3
SUP05_cluster;120
16s Copies/L/µm
Particle Size (µm)
0.2
1.2
5
20
53
180
500
Station
3.3
4.3

1619

1620 **C**



1621

1622

Figure S15. Amplicon sequence variants that comprised at least 1% of the total free-living (0.2 – 1.2 µm size fraction) community at at least one station in (**A**) the surface (**B**) oxycline and (**C**) the bottom depths. Only station and depth combinations where the free-living bacteria were sequenced are shown. All of these ASVs were found to be present in one or more of the larger size fractions.

Supplemental Tables

Table S1. Microbial abundances of each ASV in each location, depth and size fraction. Data have been processed to normalize sequence reads by spike concentrations and DNA in the environment to produce normalized abundances. Column names as follows

ASV: Numbered amplicon sequence variant

Reads: Number of reads of that ASV

ID: Sample ID of form station-depth-size fraction. For instance, 3-1-B-0-2 is station CB3.1, Bottom, Size fraction 0.2-1.2 µm.

SpikeReads: Number of reads that taxonomically mapped to the spike-in sequence in each sample.

conversionMultiplier: DNA per liter * 10^5 spike in copies / SpikeReads, used for calculating concentrations.

RA: Relative abundance of the sequence. A value between 0 and 1.

copiesPerL: The estimated number of each ASV per L of water in each size-fraction at each sample. Calculated by multiplying *Reads* by *conversionMultiplier*

We note that normalized copies per liter can be calculated by dividing *copiesPerL* by *Bin_Size*, and copies per mg of particles can be calculated by dividing *copiesPerL* by *MassperLiter.* This table is 6,184,875 lines long and takes 1.4GB of space uncompressed and so is uploaded as its own gzip compressed csv file.

Table S2 Taxonomic information about each amplicon sequence variant.

Kingdom:Genus – Taxonomy information about the ASV as returned by DADA2's built in classifier and the silva_nr_v132 database.

nASV: Just the number from the ASV column

TagLevel: The finest taxonomic level for which taxonomy is available

Tag: The taxonomic identification at TagLevel

Tag_ASV: Tag, appended to nASV. Used when naming ASVs in figures so that the ASV name provides information about taxonomy.

Table S3 Sample information table.

Station: Chesapeake bay program station ID. All stations are central bay stations (so 4.3 means CB 4.3C and so on).

Size_Class: The lower bound of the sample size class. For instance 0.2 means the 0.2 – 1.2 µm size fraction, while 500 is the ≥500 µm size fraction.

Depth: Depth category. Surface or Bottom. Station CB 4.3 also has a third Oxycline (Oxy) depth.

Bin_Size: The width of the particle size bin (µm). Calculated as the upper bound minus the lower bound.

DNAperLiter: The amount of DNA extracted from a sample, normalized to volume filtered.

MassperLiter: The amount of POM mass measured, normalized to volume filtered. Not recorded for the 0.2 – 1.2 µm size bin. This value was previously reported in Dougherty et al. (2021).

ParticlesPerLiter: Estimated number of particles in a size class, per liter. Measured by LISST and summed over all size bins. This value was previously reported in Dougherty et al. (2021).

Flag: Some samples did not amplify and/or sequence well. This is especially the case with all first run samples from station 5.1B. Only samples where FLAG is false are included here, so this is just a column of the word false.

Table S4. Relationships between metrics of alpha diversity (Richness, Shannon Diversity and Evenness) to water salinity, oxygen concentration and particle size. : **Metric:** the alpha diversity metric used. Richness (breakaway – betta) uses the betta function in the breakaway package to carry out the model. All other metrics, use a simple linear model. **Term** indicates the coefficient for which statistics are shown. Shown are estimates and standard errors of coefficients, T values (linear models only) and *p*-values (bolded if $p < 0.05$).

Metric	Term	Estimate	Std. Err.	T	*p*
Richness (Breakaway – Betta)	Intercept	4.0×10^{3}	2.4×10^{2}	NA	**< 0.001**
	$\log_{10}$(Size Class)	6.5×10^{2}	1.4×10^{2}	NA	**< 0.001**
	$\log_{10}(\text{Size Class})^2$	-4.0×10^{2}	6.2×10^{1}	NA	**< 0.001**
	Salinity	-3.7×10^{2}	2.1×10^{1}	NA	**< 0.001**
	Salinity^2	1.7×10^{1}	1.2×10^{0}	NA	**< 0.001**
	$\log_{10}$(Oxygen + 0.03)	-8.8×10^{1}	3.1×10^{2}	NA	0.779
Richness (Breakaway – LM)	Intercept	3.9×10^{3}	7.7×10^{2}	5.059	**< 0.001**
	$\log_{10}$(Size Class)	7.5×10^{2}	4.4×10^{2}	1.688	0.096
	$\log_{10}(\text{Size Class})^2$	-4.4×10^{2}	1.9×10^{2}	-2.348	**0.022**
	Salinity	-4.5×10^{2}	1.7×10^{2}	-2.610	**0.011**
	Salinity^2	2.4×10^{1}	9.1×10^{0}	2.586	**0.012**
	$\log_{10}$(Oxygen + 0.03)	4.4×10^{2}	7.6×10^{2}	0.574	0.568
Diversity (Shannon H)	Intercept	5.2×10^{0}	2.0×10^{-1}	26.613	**< 0.001**
	$\log_{10}$(Size Class)	5.0×10^{-1}	1.1×10^{-1}	4.461	**< 0.001**
	$\log_{10}(\text{Size Class})^2$	-2.1×10^{-1}	4.8×10^{-2}	-4.278	**< 0.001**
	Salinity	-9.5×10^{-2}	4.4×10^{-2}	-2.175	**0.033**
	Salinity^2	3.2×10^{-3}	2.3×10^{-3}	1.386	0.170

	$\log_{10}$(Oxygen + 0.03)	-1.4×10^{-1}	1.9×10^{-1}	-0.738	0.463
Evenness (Pielou J)	Intercept	6.1×10^{-4}	6.6×10^{-3}	0.092	0.927
	$\log_{10}$(Size Class)	-8.3×10^{-3}	3.8×10^{-3}	-2.182	**0.033**
	$\log_{10}$(Size Class)^2	3.2×10^{-3}	1.6×10^{-3}	1.965	0.053
	Salinity	1.0×10^{-3}	1.5×10^{-3}	0.687	0.494
	Salinity^2	-3.4×10^{-5}	7.8×10^{-5}	-0.432	0.667
	$\log_{10}$(Oxygen + 0.03)	4.8×10^{-3}	6.5×10^{-3}	0.738	0.463

1685

1686

Table S5. Coefficients of an ANOVA like permutation test of the RDA analysis coefficients describing the relationship between water salinity, oxygen concentration and particle size and overall microbial community structure. Three Datasets are run: All (Figure S11), which includes all size classes. Particle associated (Figure S12A) which exclude the free-living 0.2 – 1.2 µm size class samples and ≥ 5 Micron Particles which exclude both the 0.2– 1.2 µm and 1.2 – 5 µm size classes. **Term** indicates the coefficient for which statistics are shown. **DF** corresponds to degrees of freedom. **Variance** is the total variance explained by each coefficient. **% Variance** is the variance divided by total variance. **T** is the relevant statistic for the test and ***p*** is the p-value of that statistic**.**

Dataset	**Term**	**DF**	**Variance**	**% Variance**	**T**	***p***
All	Salinity	1	745	6.9%	7.0	< **0.001**
	$\log_{10}$(Oxygen + 0.03)	1	458	4.2%	4.3	< **0.001**
	$\log_{10}$(Size Class)	1	2002	18.5%	18.7	< **0.001**
	Residual	71	7613	70.4%		
Particle Associated	Salinity	1	781	8.2%	8.7	< **0.001**
	$\log_{10}$(Oxygen + 0.03)	1	546	5.7%	6.1	< **0.001**
	$\log_{10}$(Size Class)	1	2616	27.5%	29.1	< **0.001**
	Residual	62	5567	58.5%		
≥ 5 Micron Particles	Salinity	1	534	9.5%	8.2	< **0.001**
	$\log_{10}$(Oxygen + 0.03)	1	432	7.7%	6.6	< **0.001**
	$\log_{10}$(Size Class)	1	1229	21.8%	18.9	< **0.001**
	Residual	53	3445	61.1%		