Microbial Ecology (2021) 81:601–616
https://doi.org/10.1007/s00248-020-01621-7

MICROBIOLOGY OF AQUATIC SYSTEMS

In-depth Spatiotemporal Characterization of Planktonic Archaeal and Bacterial Communities in North and South San Francisco Bay

Anna N. Rasmussen[1] · Julian Damashek[1,2] · Emiley A. Eloe-Fadrosh[3] · Christopher A. Francis[1]

Received: 1 July 2020 / Accepted: 12 October 2020 / Published online: 5 November 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Despite being the largest estuary on the west coast of North America, no in-depth survey of microbial communities in San Francisco Bay (SFB) waters currently exists. In this study, we analyze bacterioplankton and archaeoplankton communities at several taxonomic levels and spatial extents (i.e., North versus South Bay) to reveal patterns in alpha and beta diversity. We assess communities using high-throughput sequencing of the 16S rRNA gene in 177 water column samples collected along a 150-km transect over a 2-year monthly time-series. In North Bay, the microbial community is strongly structured by spatial salinity changes while in South Bay seasonal variations dominate community dynamics. Along the steep salinity gradient in North Bay, we find that operational taxonomic units (OTUs; 97% identity) have higher site specificity than at coarser taxonomic levels and turnover ("species" replacement) is high, revealing a distinct brackish community (in oligo-, meso-, and polyhaline samples) from fresh and marine end-members. At coarser taxonomic levels (e.g., phylum, class), taxa are broadly distributed across salinity zones (i.e., present/abundant in a large number of samples) and brackish communities appear to be a mix of fresh and marine communities. We also observe variations in brackish communities between samples with similar salinities, likely related to differences in water residence times between North and South Bay. Throughout SFB, suspended particulate matter is positively correlated with richness and influences changes in beta diversity. Within several abundant groups, including the SAR11 clade (comprising up to 30% of reads in a sample), OTUs appear to be specialized to a specific salinity range. Some other organisms also showed pronounced seasonal abundance, including *Synechococcus*, *Ca*. Actinomarina, and *Nitrosopumilus*-like OTUs. Overall, this study represents the first in-depth spatiotemporal survey of SFB microbial communities and provides insight into how planktonic microorganisms have specialized to different niches along the salinity gradient.

Keywords Estuary · Microbial ecology · 16S rRNA · Bacterioplankton · Archaeoplankton

Introduction

San Francisco Bay (SFB) is the largest estuary on the west coast of the continental United States and is surrounded by approximately 7.8 million people (US Census Bureau 2018). Intense urban development along the shores and other human activities such as damming, diking, dredging, historic mining, and pollution have led SFB to be considered one of the most anthropogenically altered estuaries in the USA [1]. Long-term monitoring projects from both federal and state agencies have also made SFB one of the most extensively studied estuaries in the world [2], and consequently, SFB has served as a model for understanding physical, chemical, and biological estuarine dynamics [3–6]. For five decades, water quality in SFB has been monitored regularly by the United States Geological Survey (USGS) [7], showing both gradual and abrupt changes in water quality due to human activity [8–10] and leading to a thorough characterization of phytoplankton dynamics [11–14]. However, in contrast to the in-depth monitoring of phytoplankton in SFB, bacterioplankton and archaeoplankton populations are remarkably understudied in this system. In

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00248-020-01621-7.

✉ Christopher A. Francis
caf@stanford.edu

[1] Department of Earth System Science, Stanford University, 473 Via Ortega, Y2E2 Bldg Rm 140, Stanford, CA 94305, USA

[2] Present address: Department of Biology, Utica College, Utica, NY 13502, USA

[3] Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

fact, only three studies to date have examined pelagic microbial community structure in SFB [15–17], all of which exclude archaea and used molecular approaches with limited phylogenetic resolution (i.e., DGGE, T-RFLP) or sequencing depth. Here, we build considerably on this literature, using deep 16S rRNA gene amplicon sequencing at large spatial and temporal scales to understand bacterial and archaeal ecology in the turbid estuarine waters of SFB and compare the two connected but distinct arms of the estuary.

SFB is an excellent model system for understanding microbial community dynamics because it encompasses both spatial (e.g., salinity) and temporal (e.g., temperature) physicochemical gradients [10]. Salinity has been identified as a universal driver of bacterial and archaeal community structure [18–20], as well as a key driver of estuarine bacterial community composition in numerous studies [21–27]. However, studies in estuarine environments have also identified drivers of microbial community structure besides salinity, including temperature, pH, dissolved oxygen, water residence time, organic carbon, or nutrient availability [15, 22, 26, 28–30], factors that have not yet been studied in this system. Here we use the strong estuarine gradient in North Bay to understand how communities change along a salinity gradient and also compare samples from similar salinities but in two distinct subestuaries of SFB (i.e., North versus South Bay).

We sampled 12 stations ranging from fresh riverine inputs to brackish mixing zones to highly marine-influenced regions along a ~ 150-km transect of the SFB channel in order to assess microbial biogeography. We sequenced bottom water samples collected monthly over 2 years, capturing microbial communities across several seasonal gradients (e.g., temperature, freshwater flow rate). Amplicon libraries were generated using two updated 16S rRNA gene primer sets, which amplify variable region 4 [31, 32] and regions 4 and 5 [32], respectively. In our study, we use this large dataset to compare North and South Bay microbial community assembly. We calculate diversity metrics across different taxonomic levels, spatial extents, and physicochemical extents to better understand how environmental gradients influence estuarine microbial communities.

Materials and Methods

Study Site

SFB consists of two distinct but connected arms, generally referred to as North and South Bay. North Bay receives freshwater inputs from the Sacramento and San Joaquin Rivers and is characterized by a large salinity gradient ranging from fresh to fully marine with residence times varying from days to months depending on freshwater flow [33]. South Bay is a weakly mixed marine lagoon with freshwater inputs dominated by urban runoff and wastewater and residences times ranging from weeks to several months in the summer [33]. North and South Bay consist of several sub-embayments (labeled in Fig. 1) and are connected by a deeper basin, Central Bay, which is strongly influenced by tidal exchange with the Pacific Ocean [3].

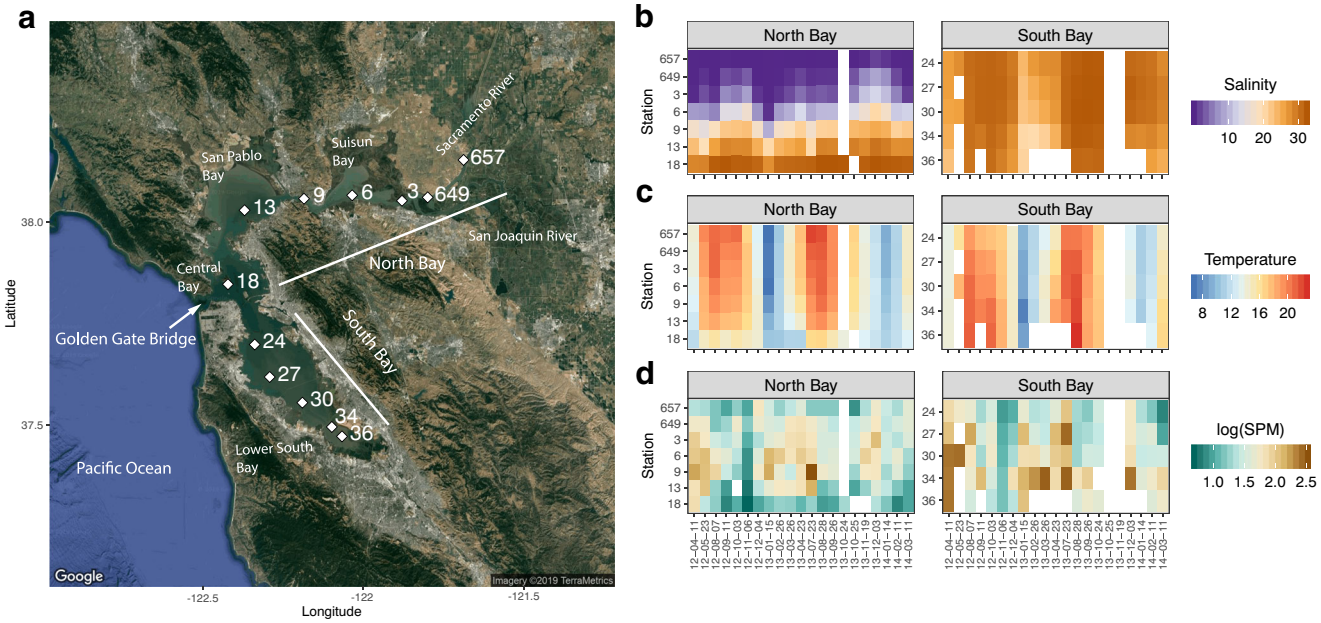

Fig. 1 Map of San Francisco Bay and USGS water quality stations sampled in this study (**a**) and tile plots of salinity in PSU (**b**), temperature in °C (**c**), and SPM in mg/L (**d**) with the *y*-axis corresponding to stations and *x*-axis to dates (YY-MM-DD) sampled in this study. Plots are faceted by North (left facet) and South (right facet) Bay

Sampling and DNA extraction

Microbial biomass was collected for DNA extraction approximately monthly between April 2012 and March 2014 during United States Geological Survey (USGS) Water Quality monitoring cruises in the channel of the San Francisco Bay estuary. Microbial cells were collected from bottom waters (1-m above estuary floor) by pressure-filtering 150–1000 mL of water from Conductivity, Temperature, and Depth (CTD) instrument package casts through a 10-μm pore size polycarbonate Isopore membrane filter (47-mm diameter; EMD Millipore, Darmstadt, Germany) in line with a 0.22-μm polyethersulfone Supor-200 membrane filter (47-mm diameter; Pall, Port Washington, NY), followed by flash freezing on liquid nitrogen prior to storage at − 80 °C. Bottom waters were the focus of this study because while phytoplankton in surface waters have been well studied, very little is known about microbial processes/communities below the surface in SFB. Due to the turbid waters and rapid clogging of small pore-sized filters, a 10-μm prefilter was used to increase bacterial and archaeal biomass collection. Only 0.22-μm filters were frozen at − 80 °C and saved for extraction. DNA was extracted with the FastDNA SPIN Kit for Soil (MP Biomedicals, Santa Ana, CA), following the manufacturer's instructions, with the following modifications: bead tubes were homogenized for 40 s at speed 6.0 in a FastPrep bead beater (MP Biomedicals, Santa Ana, CA), and final DNA was eluted into 75 μL 55 °C sterile DNase-free water. DNA was quantified using the Qubit dsDNA Broad Range assay (Life Technologies, Grand Island, NY) and stored at − 80 °C.

Environmental Data

Corresponding water quality data [7] from the exact sampling cruises, stations, and bottom water depth used in this study was downloaded from the USGS San Francisco Bay Water Quality and California Day Flow websites, the data and detailed methodology on measurements are available at the following links: https://sfbay.wr.usgs.gov/access/wqdata/query/index.html and http://www.water.ca.gov/dayflow/. Suspended particulate matter (SPM) was measured based on optical backscatter and corrected with discrete measurements of dry weight of filtered particles. Additional ammonium, nitrate, and nitrite measurements were performed using filtered (0.22 μm pore size) water that was frozen on dry ice prior to storage at − 20 °C. Ammonium was measured using the salicylate-hypochlorite method (Bower and Holm-Hansen, 1980). Nitrate and nitrite were measured using a SmartChem200 Discrete Analyzer (Unity Scientific, Brookfield, CT) following standard procedures. Nutrients were measured within 1 week of sample collection.

Sequencing

Through a DOE Joint Genome Institute (JGI) Community Science Program (CSP) project, we have assembled 354 16S rRNA gene Illumina (MiSeq) amplicon libraries from bottom water samples collected on 20 approximately monthly cruises at 12 USGS monitoring stations spanning a 150-km transect (Fig. 1). Of the 180 samples submitted, 177 samples were successfully amplified with the updated 16S V4 primers [31, 34] (515F-Y GTGYCAGCMGCCGCGGTAA and 806RB GGACTACNVGGGTWTCTAAT) and 177 samples were amplified with the 16S V4-V5 primers [32] (515F-Y and 926R CCGYCAATTYMTTTRAGTTT). In total, 174 samples were amplified by both primer pairs to ensure we captured as much of the microbial diversity as possible (particularly in the SAR11 clade and *Thaumarchaeota*, which previous V4 primers were biased against [31, 32, 34]) and allow us to compare how these "universal" sets described an estuarine environmental dataset. Library preparation was performed by JGI following their standard operating procedures, available at: https://jgi.doe.gov/user-programs/pmo-overview/protocols-sample-preparation-information/. Libraries were amplified at 94 °C for 3 min, then 30 cycles of 94 °C for 45 s, 50 °C for 60 s, 72 °C for 90 s, with a final elongation of 72 °C for 10 min before holding indefinitely at 4 °C. Samples were pooled and sequenced on an Illumina MiSeq. Sequence data from this study is available on NCBI SRA under BioProject PRJNA577706.

Data Processing

Raw reads were processed by JGI using the iTagger v2.2 method, which is described in detail at https://bitbucket.org/berkeleylab/jgi_itagger. Briefly, this method used Usearch v9.2.64_i86linux32 [35], MAFFT v7.310 (2017/Mar/17) [36], and QIIME 1.9.1 [37]. Using Usearch, overlapping read pairs are merged into unpaired consensus sequences; merged reads are then checked for correct PCR primer orientation and spacing, evaluated based on read quality scores, dereplicated, and counted. Next, clusterable sequences were incrementally clustered starting at 99% identity and increasing the radius by 1% until reaching 97% to define OTUs. Sequences with fewer than 3 copies were not used to cluster and were later mapped back to cluster centroids. Cluster centroid sequences were evaluated with the reference database SILVA 128. OTUs with no taxonomic classification were moved to an otu.unknown.fasta file, and contaminating sequences (i.e., chloroplast) were removed. After clustering, the V4 dataset contained 160,616,905 reads (91.7% of original reads) with an average library size ~ 890,000 and the V4–V5 dataset contained 111,611,495 reads (91.9% of original reads) with an average library size ~ 610,000.

In general, amplicon library processing followed established protocols for filtering low abundance reads, transforming read counts, and normalizing library sizes [38, 39]. To prevent high-variance-low-abundance OTUs from strongly influencing downstream analyses, we removed OTUs that did not have at least 16 reads in at least 2 samples. Libraries were normalized using the *DESeq2* [40] variance-stabilizing transformation (VST) method to adjust counts based on variation in library size and then transformed using the geometric mean to account for large variation in OTU counts [39]. All analyses were conducted with R [v3.5.1] using primarily *phyloseq* v1.26.1 [38, 39, 41] and *vegan* v2.5-4 [42] packages.

Environmental data was centered and scaled prior to correlation testing (*scale*, *cor.test*, and *ggpairs* [*GGally v1.4.0*] functions). The natural log was used to transform variables where values spanned several orders of magnitude (e.g., suspended particulate matter, chlorophyll). Samples were assigned into salinity zones defined as fresh, oligohaline, mesohaline, polyhaline, and euhaline [Venice system, [43]]. The wet season was defined as December through May and the dry season as June through November, based on flow rates from the Delta.

Community Diversity Analysis

Observed OTUs, Chao1, and Inverse Simpson metrics were calculated for unfiltered (i.e., contain singletons) amplicon libraries using *phyloseq*. Beta diversity metrics (e.g., principal coordinates analysis, distance-based redundancy analysis) were based on Bray-Curtis dissimilarity unless otherwise noted. For beta diversity analyses within salinity zones, we used relative abundance transformed libraries. *DESeq2* was used for differential abundance testing between different salinity zones. The *clusGap* function [*cluster* v2.0.7-1] was used to calculate the gap statistic for "goodness of clusters" and evaluated using "TibsSE2001SEmax" and used to confirm that the number of clusters in *adonis* tests was appropriate. The *bioenv* [*vegan*] function was used to identify the environmental variables with the best correlation with community dissimilarities from the following 10 variables: salinity, depth, temperature, chlorophyll *a*, chlorophyll *a* to phaeopigments ratio, nitrate, nitrite, ammonium, SPM, and freshwater flow rate. In *vegan*, analysis of variance was calculated using the *anova.cca* or *adonis* function and dispersions were calculated and tested using *betasdisper* and *permutest*. Significant variables identified using *bioenv* test were used in dbRDAs unless a model with fewer environmental variables could explain more variance based on *anova* tests. Procrustes tests were used to compare ordinations quantitatively using the *protest* function [*vegan*], which rotates and stretches ordinations until the distance between corresponding points is minimized. M^2 represents one minus the squared correlation coefficient (R) between the coordinates of corresponding points between the two ordinations being tested. Turnover versus nestedness components of beta diversity analyses were calculated using *betapart* v1.5.1.

Nestedness of samples was calculated using the NODF statistic (Nestedness metric based on Overlap and Decreasing Fill), which describes combined column and row nestedness based on decreasing fill and paired overlap of presence data in a matrix [44]. To calculate NODF, count tables were converted to a presence-absence matrix and ordered by decreasing row sums, then decreasing column sums (decreasing prevalence and decreasing richness, respectively). The *oecosimu* function [*vegan*] was used to calculate the NODF statistic, which approximates the average percent of taxa from less diverse samples that occur in more diverse samples. The null model used was the "c0" model, which holds constant the column sums (sample richness) but allows row sums to vary (phyla prevalence) from the original data. Significance testing was based on default parameters for *oecosimu* package and set to test if data values were greater than the null model. Entropy was calculated using the *diversity* function [*vegan*] and defined as the Shannon index of specific taxa (instead of by sample) based on natural log of relative abundance.

Results and Discussion

Primer Comparison

For the remainder of this text, all analyses presented will be based on the 16S rRNA gene libraries amplified by the 515F-Y [32] and 806RB [31] primers. A comparison of this "V4 dataset" and libraries amplified by 515F-Y and 926R [32] "V4–V5" dataset revealed strikingly similar alpha and beta diversity metrics (Supplemental Fig. S1). The V4 and V4–V5 primers also similarly described SAR11 clade bacteria and *Thaumarchaeota* distributions along the salinity gradient (data not shown), in general agreement with rigorous comparisons of these primers made in the marine environment [45].

North Bay

Environmental Data

Water column samples within this dataset correspond to salinities ranging from fresh (minimum 0.07 practical salinity units [PSU]) to euhaline (maximum 32.42 PSU) and temperatures from 6.8 to 22.1 °C (Fig. 1). Stations typically falling into the Venice salinity zones [43] are as follows: station 657 was fresh (< 0.5 PSU), station 649 is oligohaline (0.5 to < 5 PSU), station 6 was mesohaline (5 to < 18 PSU), station 13 was polyhaline (18 to < 30 PSU), and station 18 was euhaline (30 to < 40 PSU) (Supplemental Fig. S2). Chlorophyll *a*

concentrations were typically low (median = 2.5 μg/L) with peaks concentrations occurring in March (Supplemental Fig. S3). Ammonium concentrations were highest in riverine samples (Supplemental Fig. S3) due to inputs from the Sacramento Regional Wastewater Treatment Plant [46]. Agricultural runoff into the Delta can be a major source of nutrients, particularly nitrate, in North SFB [46].

Salinity Dominates Changes in Beta Diversity

Salinity is a key factor shaping microbial communities at a global scale [18–20] and in a wide variety of estuaries [22, 25, 26, 47, 48]. We used several different metrics to assess beta diversity across the strong salinity gradient in North Bay. First, we used ordination analyses to assess beta diversity at the OTU level. An unconstrained PCoA shows that 65.9% of community dissimilarity is explained by the first three axes (Supplemental Fig. S4). Salinity is strongly correlated with PC1 (r^2 = 0.92), PC2 is most strongly correlated with SPM (r^2 = 0.40), and PC3 is most strongly correlated with temperature (r^2 = 0.50). Given the strong structure based on salinity, we assessed the relevance of salinity zones for defining clusters. Centroids of salinity zones are significantly different (adonis, $p < 0.001$, Supplemental Fig. S4); however, beta dispersions are also significantly different between the three "brackish" zones (oligo-, meso-, and polyhaline) and euhaline and fresh samples (permutest, $p = 0.001$). This supports that different brackish (oligo-, meso-, and polyhaline) communities are distinct from fresh and marine end-members but also distinct from each other. We were also interested in understanding clusters observed in PC2 and PC3 and found that samples cluster significantly into slightly overlapping wet and dry season groups (adonis, $p = 0.001$, Supplemental Fig. S4). We then sought to understand how variation in communities could be explained by changes in environmental factors. Salinity, temperature, chlorophyll *a*, and ratio of chlorophyll *a* to phaeopigments have the highest rank correlation with the Bray-Curtis dissimilarity matrix (bioenv, $\rho = 0.55$). A dbRDA using these same factors was also performed but we found that a constrained dbRDA using only 3 environmental factors (salinity, temperature, and SPM) could explain more community variation (52.2% versus 55.2%, respectively) (Fig. 2; adonis, $p < .001$). Salinity has by far the strongest influence on community structure in the dbRDA with a marginal effect of 41.0% (ANOVA, $p = 0.001$), followed by SPM (8.2%, ANOVA, $p = 0.001$) and temperature (6.2%, ANOVA, $p = 0.001$). Both constrained and unconstrained ordinations reveal a strong influence of salinity, followed by changes in SPM, and seasonal changes such as temperature, on community dissimilarity (Fig. 2 and Supplemental Fig. S4). While it is perhaps not surprising that community structure was so strongly tied with salinity along the gradient, these analyses highlight the unique "brackish" communities found at intermediate salinities. However, these analyses also highlight the importance of SPM dynamics, which will be discussed in further detail in a later section.

Salinity Tolerance Varies at Different Taxonomic Levels

To further understand how communities change along the strong salinity gradient in North Bay, we looked more deeply at the two components of beta diversity, nestedness (i.e., "species" loss) and turnover (i.e., "species" replacement)[44], and taxa entropy (i.e., site specificity) [19]. Here "species" means a taxonomic group and "site" means sample. We used the NODF statistic (approximate average percent of taxa from less diverse samples occurring in more diverse samples) to calculate nestedness. We find that communities are highly nested at coarser taxonomic levels (e.g., phylum, class) and become substantially less nested and not significantly different from

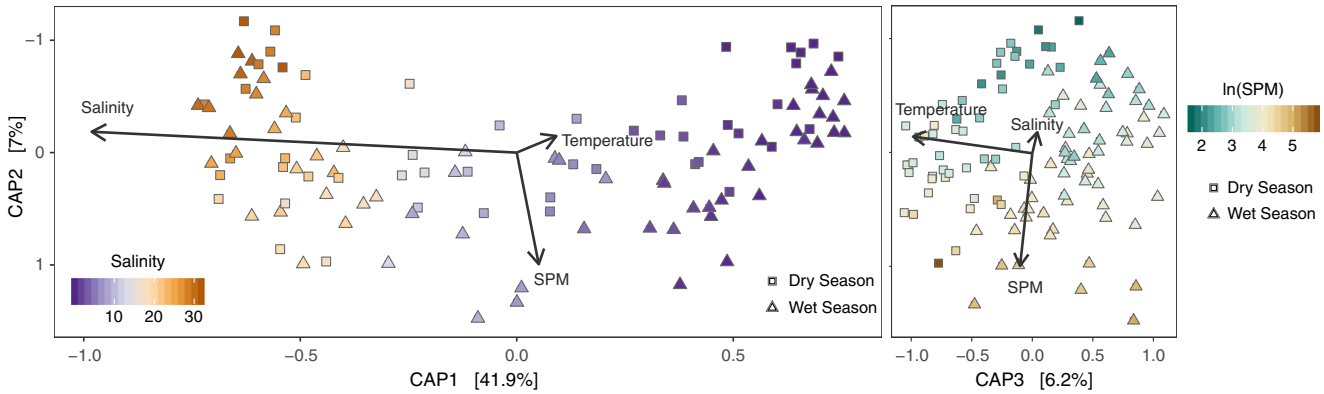

Fig. 2 North Bay beta diversity shown through dbRDA plots based on Bray-Curtis dissimilarity and three environmental factors. The first three axes are shown and point shape corresponds to season (circle for dry, square for wet). The wet season is defined as the months December through May based on increased flow rates due to rain and snowmelt; the dry season is defined as June through November based on lower flow rates during sampling years. Each CAP axis is a constrained ordination axis based on a linear combination of the environmental variables used in the model

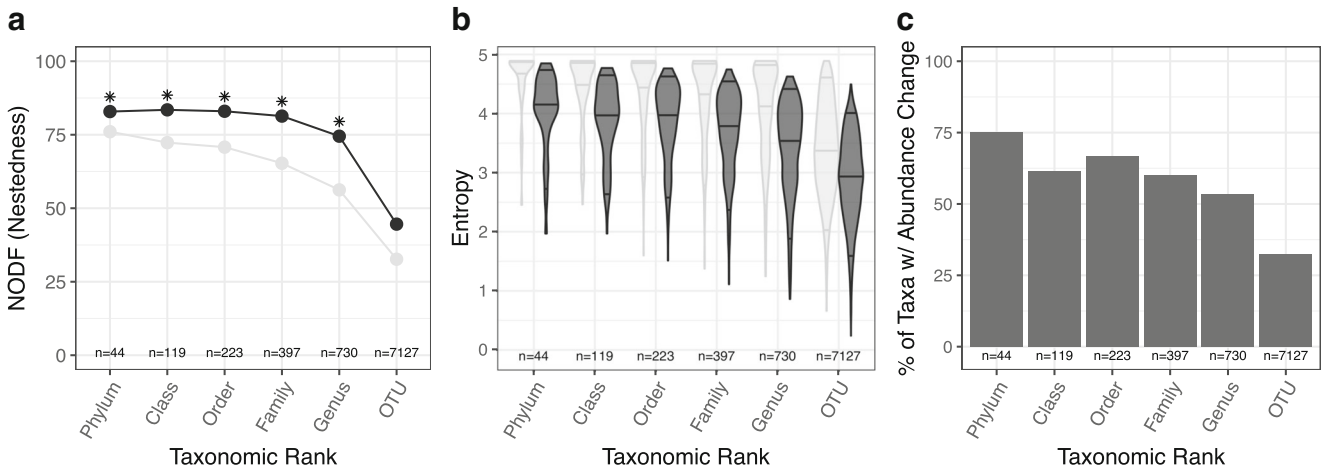

Fig. 3 Diversity metrics for North Bay, including the NODF statistic at different taxonomic levels (**a**)as a measure of nestedness for each sample and compared to a null model with constant column sums (sample richness) but changing rows (taxa prevalence). *A value is greater than null NODF with a $p < 0.05$. **b** Entropy calculated as the Shannon index of taxa. Violin plots are scaled to have the same width and horizontal lines indicate the 50th quartile based on point distribution. **c** The % of taxa with significant (p value < 0.01) log2-fold changes in abundance across the 5 salinity zones calculated using *DESeq2*

the null at the finest taxonomic level examined, OTUs (Fig. 3). We also used the *betapart* package to confirm that turnover ("species" replacement [49], beta.SIM = 0.96) is a much larger component than nestedness ("species" loss [49], beta.SNE = 0.0084) in community dissimilarity metrics (beta.SOR = 0.97). These findings highlight that "species" loss is an important process at coarser taxonomic levels, while "species" replacement is important at the OTU level.

We used the Shannon index to calculate the entropy of taxa, which we used as a measure of site specificity. Taxa can be present or abundant in many samples (high entropy/low site specificity) or a few samples (low entropy/high site specificity). Entropy of taxa is high at the coarsest taxonomic level (i.e., phylum) with a median entropy of 4.14 and decreases with finer taxonomic resolution, particularly at the OTU level, which has a median entropy of 2.94 (Fig. 3), indicating that OTUs have high site specificity. Examples of taxa with high entropy include phyla such as *Bacteroidetes* (4.85), *Proteobacteria* (4.84), and *Actinobacteria* (4.83), orders (see Fig. 4) such as *Flavobacterales* (4.76) and SAR11 clade bacteria (4.72), and many highly abundant OTUs including OTUs 6, 14, 21, 39, and 56 (Fig. 4, Supplemental Table S2). Taxa with lower entropy (high site specificity) tend to occur in or have high abundance in a small number of samples; examples include phyla such as *Thermotogae* (2.71), *Gracilibacteria* (2.51), and *Cloacimonetes* (1.97), or the order *Vibrionales* (2.86), which is low abundance in most samples (< 0.2 %) except during October 2013 when it reaches up to 10% abundance at Station 649. To further understand the distribution of organisms across samples, we also used differential abundance testing to understand how taxa abundance varies between salinity zones. Differential abundance testing shows that many taxa (from 30% at the OTU level to 75% at the phylum level) across all taxonomic levels have significant changes in abundance between the five salinity zones (Fig. 3). Thus, while many phyla are present in samples across the salinity gradient (high entropy), the abundance of most phyla varies significantly, supporting that there are still ecologically meaningful adaptations based on salinity even at the coarse phylum level.

Patterns in community nestedness, entropy, and abundance support that at the phylum level organisms have broader salinity tolerance (i.e., can be found in samples across a wide range of salinities) but organisms at the 97% identity level are adapted to a specific salinity range. Taken together with the large turnover (i.e., "species" replacement) component of beta diversity metrics, our results support that turnover is high along the salinity gradient and that related organisms replace their more or less salt-tolerant relatives along the gradient. Bar plots reveal clear patterns in each salinity zone (Fig. 4), with a gradual distinction between freshwater and euhaline sites at the class and order levels, followed by more specific mesohaline communities emerging at the genus and OTU levels. The most abundant organisms in our study at the phylum (i.e., *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*) and class (i.e., *Alpha-*, *Beta-*, and *Gammaproteobacteria*) levels have frequent transitions between marine and freshwater environments [20]. Our study strongly supports the existence of distinct brackish bacterial communities and that closely related organisms specialize/adapt to distinct salinity regimes [20, 26, 50, 51]; furthermore, it also highlights how adaptation to intermediate salinities could allow for transitions between marine and fresh environments. Interestingly, metagenomic analysis of the Baltic Sea found a distinct brackish microbiome comprised of organisms containing streamlined genomes [52]. These organisms were distinct from those of freshwater and marine

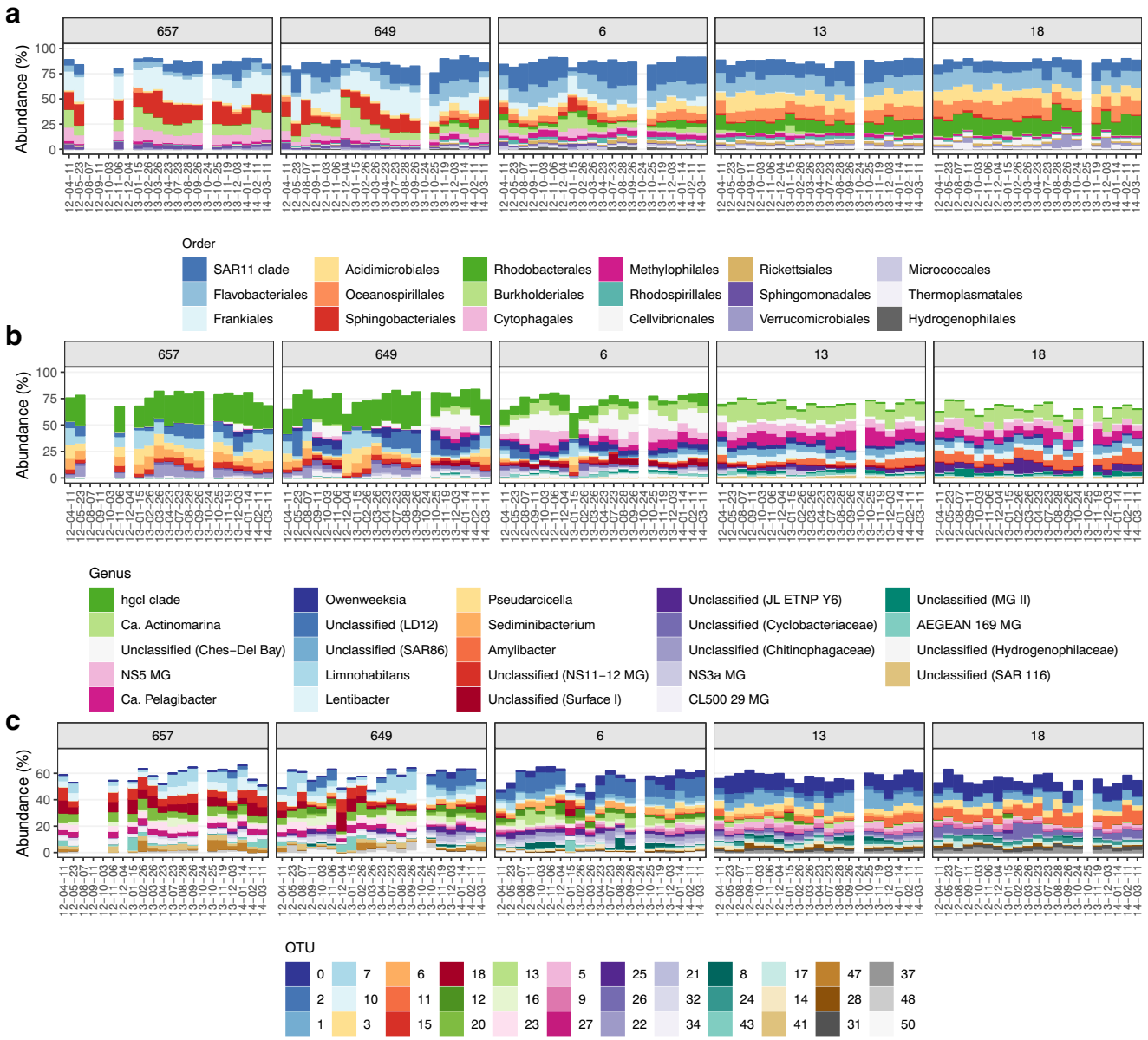

Fig. 4 The % relative abundance of the 16 most abundant orders (**a**), 20 most abundant genera (**b**), and 36 most abundant OTUs (**c**) at 5 stations representative for the five salinity zones in North Bay. The *y*-axis shows the relative abundance of taxa in the whole community (note the difference in scale in **c**), and the *x*-axis of each panel corresponds to the date (YY-MM-DD) of 20 approximately monthly cruises from April 2012 to March 2014, with the station indicated at the top of the panel. The colors do not correspond between a, b, and c

systems, and closely related organisms were found within intermediate salinities of estuarine systems around the globe.

Biogeography of Abundant Microbial Taxa

In North Bay, *Proteobacteria* is the most abundant phylum and is widely distributed along the gradient. Within the *Proteobacteria*, we observe decreasing *Betaproteobacteria* and increasing *Alpha-* and *Gammaproteobacteria* with increasing salinity, as observed in other estuaries [16, 21, 26, 53–55]. At the order level, *Oceanospirillales* and *Rhodobacterales* are more abundant in marine-influenced samples whereas *Burkholderiales* dominates in freshwater end-member stations (Fig. 4). SAR11 clade bacteria (*Alphaproteobacteria*) are abundant along the entire salinity gradient and are some of the most abundant organisms at the order level, genus level (i.e., *Pelagibacter*), and OTU level (OTUs 2, 1, 7, and 22; Figs. 4 and 5). Previous studies in SFB found SAR11 clade bacteria to be ubiquitous [16] and freshwater SAR11 (LD12-like organisms) to occur solely in fresh sites [17]. In our study, we find that SAR11 clades are very abundant and highly structured along the salinity

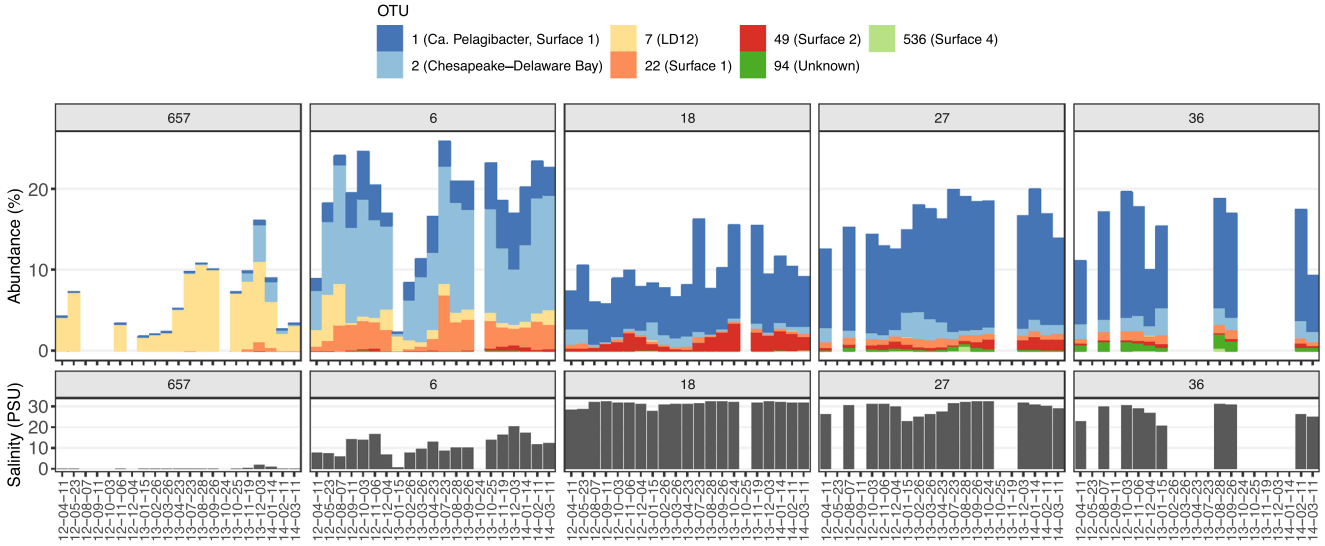

Fig. 5 The % relative abundance of SAR11 OTUs (with family indicated in parentheses) over 0.2% abundance throughout SFB. The *y*-axis shows the relative abundance of taxa in the whole community, and the *x*-axis of each panel corresponds to the date (YY-MM-DD) of 20 approximately monthly cruises from April 2012 to March 2014, with the station indicated at the top of the panel. Below the SAR11 abundance is a bar plot indicating the salinity of a given sample

gradient in North Bay. Interestingly, the distribution of LD12-like organisms in our study aligns well with recent findings from the first cultured isolate from this low-salinity group [56], with peak LD12 abundances occurring in primarily fresh and oligohaline sites (salinity < 5 PSU; Fig. 5). LD12 are proposed to be specialized to freshwater environments through the loss of key compatible solute genes [56], highlighting one way this group may have differentiated itself from its relatives and adapted to a specific niche. A family identified as the "Chesapeake-Delaware Bay clade" (SAR11 IIIa [55]) is most abundant at mesohaline and oligohaline sites, while "Surface 1" (SAR11 Ia [57]) OTUs dominate at poly- and euhaline sites. These brackish organisms, including OTU 2 and OTU 22, are 100% identical to sequences from Chesapeake Bay (accession nos. EU802173 and EU802224 [58], respectively; Supplemental Fig. S5). While the dominance of LD12 organisms at fresh sites, Chesapeake-Delaware Bay clade at mesohaline sites, and Surface 1 and 2 in poly/euhaline sites has been observed in other estuaries [55, 59, 60], our extensive spatiotemporal dataset clearly highlights how the abundance of these groups corresponds to salinity across the gradient and provides further insight into the preferred salinity niches of key estuarine SAR11 groups.

Bacteroidetes are also abundant in North Bay, particularly *Flavobacteriia*, which are important for the breakdown of organic matter in estuaries and coasts [61, 62]. *Bacteroidetes* is the second most abundant and second richest phylum in the dataset (Supplemental Fig. S1), in agreement with previous findings that marine *Flavobacteria* have high global and local diversity [63]. *Flavobacteriales* is one of the most abundant orders in the dataset (Fig. 4) and, like SAR11, is generally broadly distributed across the salinity gradient. The "NS5 marine clade" genus of *Flavobacteriaceae* is among the top genera (Fig. 4). The uncultivated NS5 group has been associated with phytoplankton blooms or the breakdown of dissolved organic matter [64–68]. In fresh end-member stations, the families *Chitinophagaceae*, *Cytophagaceae*, and "NS11-12 marine group" [63] dominate (Fig. 4). Brackish stations are dominated by the family *Cryomorphaceae*, while marine end-member stations are dominated by *Flavobacteraceae*.

Actinobacteria is the third most abundant phylum in SFB. *Candidatus* Actinomarina and "hgcI clade" [69] are among the most abundant genera (Fig. 4), with diverse "hgcl clade" (order *Frankiales*) organisms dominating (~ 20% abundance) fresh and oligohaline stations and *Ca*. Actinomarina like OTUs (order *Acidomicrobiales*) dominating (~ 15% abundance) in poly- and euhaline sites (Fig. 4). Other estuarine studies have found high diversity and specialization of *Actinobacteria* across environmental gradients [59, 70, 71], which were observed in SFB as well. *Actinobacteria* richness is most strongly correlated to SPM ($r^2 = 0.55$), but also strongly negatively correlated with salinity ($r^2 = 0.30$). While there appears to be more OTU-level diversity in the freshwater stations, a single *Ca.* Actinomarina-like OTU (OTU 0) dominates the most marine-influenced station (18), and two OTUs (0 and 8) dominate polyhaline station 13. Both fresh and marine planktonic *Actinobacteria* have been described as photoheterotrophs, and some are capable of degradation of recalcitrant organic matter [72], highlighting their potential functional role in SFB.

Cyanobacteria are generally low abundance in our dataset; however, *Synechococcus*-like organisms reach a peak abundance of 4–6% in the late summer in mesohaline sites (Fig. 4). SFB is characterized by spring phytoplankton blooms

dominated by larger phytoplankton cells, while smaller phytoplankton dominate in summer months [73, 74], in agreement with our findings. Many coastal and estuarine studies observe peak *Synechococcus* abundances in mid or late summer [75–78]. We observe a coinciding peak abundance of *Ca.* Actinomarina OTU 8 with *Synechococcus* OTU 59 in late summer in brackish stations. Metagenomes of *Ca.* Actinomarina organisms suggest they are very small, free-living photoheterotrophic organisms with streamlined, low-GC genomes (containing novel rhodopsins), which often co-occur with *Synechococcus* in marine photic zones [79, 80].

While Archaea are generally at low abundance (0.02–0.2%) in freshwater sites of SFB, *Euryarchaeota* are seasonally abundant in poly- and euhaline sites, comprising roughly 6% of reads. *Euryarchaeota* are most abundant in summer in marine dominated sites and are primarily Marine Group II (MGII) organisms, which may be proteorhodopsin-based photoheterotrophs (e.g., degrading algal biomass) as observed in other coastal sites [81, 82]. *Thaumarchaeota* genera and OTUs are distinct at the fresh and marine end-member sites, with organisms generally considered fresh (*Nitrosarchaeum*-like) [83, 84] populating station 657 and marine (*Nitrosopelagicus*) [85] populating station 18.

South Bay

Environmental Data

South Bay samples were either poly- or euhaline (Supplemental Fig. S2) and had a much shallower salinity gradient than North Bay, with salinity ranging from 20.8 to 32.3 PSU. Chlorophyll *a* concentrations were typically low but slightly higher than in North Bay (median 4.3 μg/L) with peak concentrations occurring in February and March (Supplemental Fig. S3). Nitrate concentrations were highest in Lower South Bay due to the San Jose-Santa Clara Regional Wastewater Facility (Supplemental Fig. S3; [46]). Wastewater treatment plants are by far the dominant source of nitrate and ammonium from Central to Lower South Bay [46]. Nitrite also occasionally reached high concentrations (> 9 μM) in South Bay (Supplemental Fig. S3), consistent with previous reports of elevated nitrite in this region of the bay [86, 87].

Seasonal Variations Dominate Beta Diversity Changes in South Bay

In South Bay, PCoA ordinations reveal separate clustering of samples based on wet versus dry season (adonis, $p = 0.001$, Supplemental Fig. S6). Wet and dry season clusters also generally correspond with polyhaline and euhaline clusters, though clusters by salinity zone slightly overlap (adonis, $p = 0.001$, Supplemental Fig. S6). These patterns highlight how salinity varies strongly by season in the lagoonal South Bay; warm summers with low freshwater flow lead to long residence times and euhaline salinities, and cooler wet seasons lead to predominantly polyhaline salinities (Fig. 1 and Supplemental Fig. S2). There is also significant clustering by region, with South Bay and Lower South Bay having different centroids (adonis, $p = 0.001$, Supplemental Fig. S6). Salinity, temperature, depth, nitrate, and SPM have the highest rank correlation with community dissimilarity matrix (bioenv, $\rho = 0.71$). A dbRDA using these same factors can explain 47.1% of community variation (Fig. 6), with significant marginal effects only for SPM (13.7%, ANOVA, $p = 0.001$), salinity (4.6%, ANOVA, $p = 0.001$), temperature 3.9%, ANOVA, $p = 0.001$), and depth (3.7%, ANOVA, $p = 0.001$). These analyses highlight the complex combination of factors shaping South Bay communities, including variations in SPM, seasonal changes in salinity and temperature, as well as regional variations between South and Lower South Bay in nitrate concentrations and water column depth (Supplemental Fig. S7).

Biogeography of Abundant Microbial Taxa

In South Bay, dominant taxa include members of *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, and *Thaumarchaeota* (Fig. 7, Table S2). Because sites are either poly- or euhaline, we find abundant marine-associated organisms such as *Oceanospiralles*, *Rhodobacterales*, *Flavobacteriales*, *Ca.* Actinomarina, *Pelagibacter*, and NS5. SAR11 clade organisms in South Bay resemble those of North Bay polyhaline sites; however, there is also a SAR11 organism (OTU 94) that is more abundant in South than North Bay (Fig. 5) and is identical to environmental sequences from coastal sites such as Newport Harbor, RI (accession no. EU799850 [58]), the Lagoon of Venice (accession no. FN435239 [88]), and the Gulf of Mexico (accession no. MK603665) (Supplemental Fig. S5). There is a high abundance of *Ca.* Actinomarina (~ 20%) in South Bay, and a seasonal peak of OTU 8 at Station 27 that does not co-occur with *Synechococcus* (as was observed in North Bay), warranting further exploration of these potentially phototrophic *Actinobacteria* in SFB (Fig. 7).

We also observe a "bloom" of *Thaumarchaeota* in South Bay, with a putative *Nitrosopumilus*-like sequence (OTU 30) comprising up to 20% of the total reads, which has not been previously described for this system (Fig. 8). Blooms of *Thaumarchaeota* have been reported in other estuaries in summer [89–91], but also in coastal areas in fall [92, 93] or winter [94, 95]. This SFB bloom is associated with nitrite accumulation ($r^2 = 0.79$, Fig. 8), a phenomenon observed in other coastal and estuarine sites [90, 93]. Notably, warmer temperatures between 20 °C and 30 °C have been proposed to explain the decoupling of ammonia and nitrite oxidation in other systems [90]; however, this SFB bloom occurs in the mid to

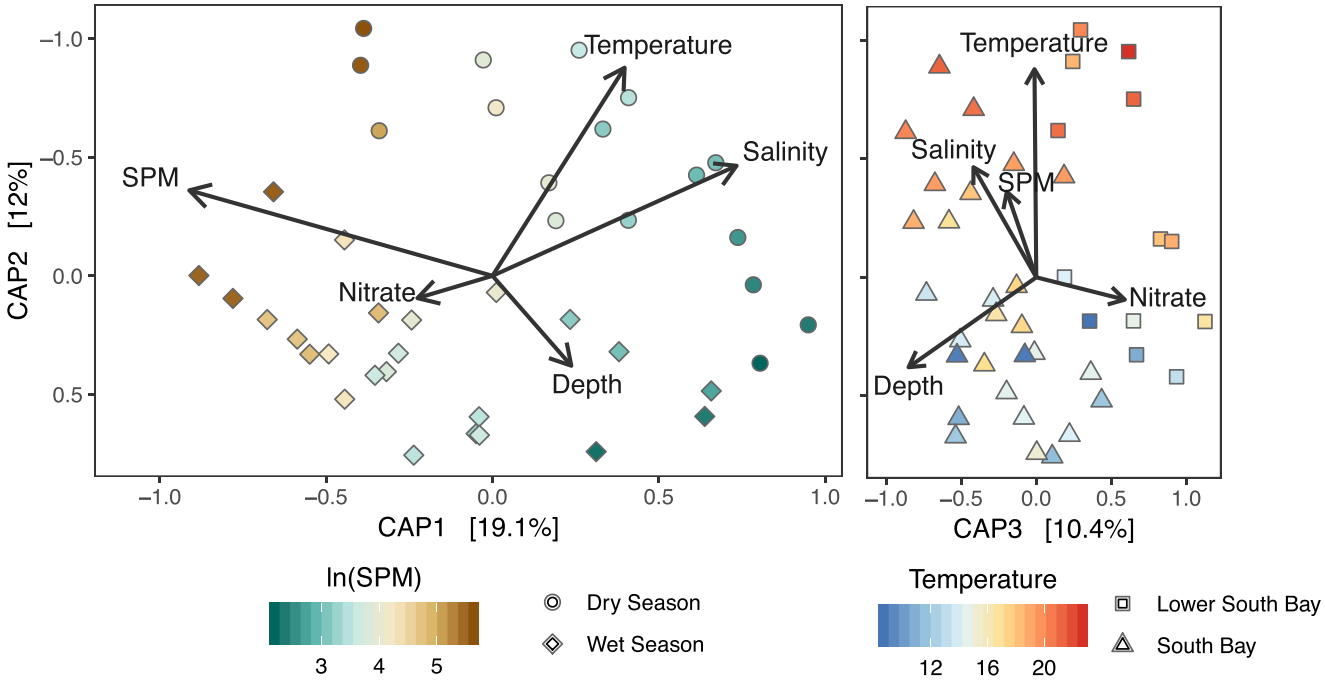

Fig. 6 South Bay beta diversity shown through dbRDA plots based on Bray-Curtis dissimilarity and four environmental factors. The first three axes are shown and point shape corresponds to season (circle for dry, square for wet). The wet season is defined as the months December through May based on increased flow rates due to rain and snowmelt; the dry season is defined as June through November based on lower flow rates during sampling years. Each CAP axis is a constrained ordination axis based on a linear combination of the environmental variables used in the model

late fall when temperatures are generally at or below 20 °C and decreasing. Nitrite accumulation can also occur in low oxygen environments, as observed after wind events stimulate ammonia but not nitrite oxidation over the stratified Chesapeake Bay water column [91]. This apparent decoupling of ammonia and nitrite oxidation in the well-oxygenated waters of South San Francisco Bay is intriguing and warrants further investigation. The high abundance of these *Thaumarchaeota* could also have significant implications for nitrogen cycling; indeed, nitrification rate measurements for this region of the bay are limited [87], but activity could be high during these blooms.

Assessing SFB as a Whole

Seasonal and Regional Variations Apparent Within Salinity Zones

Clustering in a PCoA of North Bay samples (Supplemental Fig. 4) and of all samples has significantly different centroids between all salinity zones (adonis, $p < 0.001$), supporting that salinity zone definitions are ecologically relevant for microbes. We leveraged this finding to assess diversity within salinity zones (regardless of region), since spatial/salinity gradients can overwhelm seasonal variations in estuaries [25]. PCoA ordinations reveal non-overlapping wet and dry seasonal clusters in fresh, oligohaline, and mesohaline zones, though freshwater samples also have significantly different dispersions (Supplemental Table S1). It is worth noting that because South Bay only had poly- or euhaline samples, analyses of lower salinity zones are only from North Bay samples. In ordinations of all polyhaline or all euhaline samples from the entire Bay, there are significant non-overlapping clusters based on region (North vs. South Bay, Supplemental Table S1). By assessing communities within salinity zones, we are able to reveal stronger seasonal or regional variations.

Given the strong regional differences observed in the beta diversity of poly- and euhaline samples, we wanted to further examine why microbial communities were distinct in North and South Bay despite being in the same salinity zone. To do so, we calculated community dissimilarity including poly- and euhaline samples from both North and South Bay but excluded station 18, the most tidally influenced station. In both PCoA and dbRDA plots, we note three clusters corresponding to the following: South Bay dry season samples, a mix of South Bay wet season and North Bay dry season, and North Bay wet season samples (Supplemental Fig. S8). These groupings also correspond to expected residence times on the order of months, weeks, and days, respectively [1]. Residence time is important for the formation of distinct brackish communities in other estuaries [22, 26], so it is worth noting that North and South Bay communities can be similar when residence times are expected to be of similar magnitude and they are within the same salinity zone. In addition to clusters based on

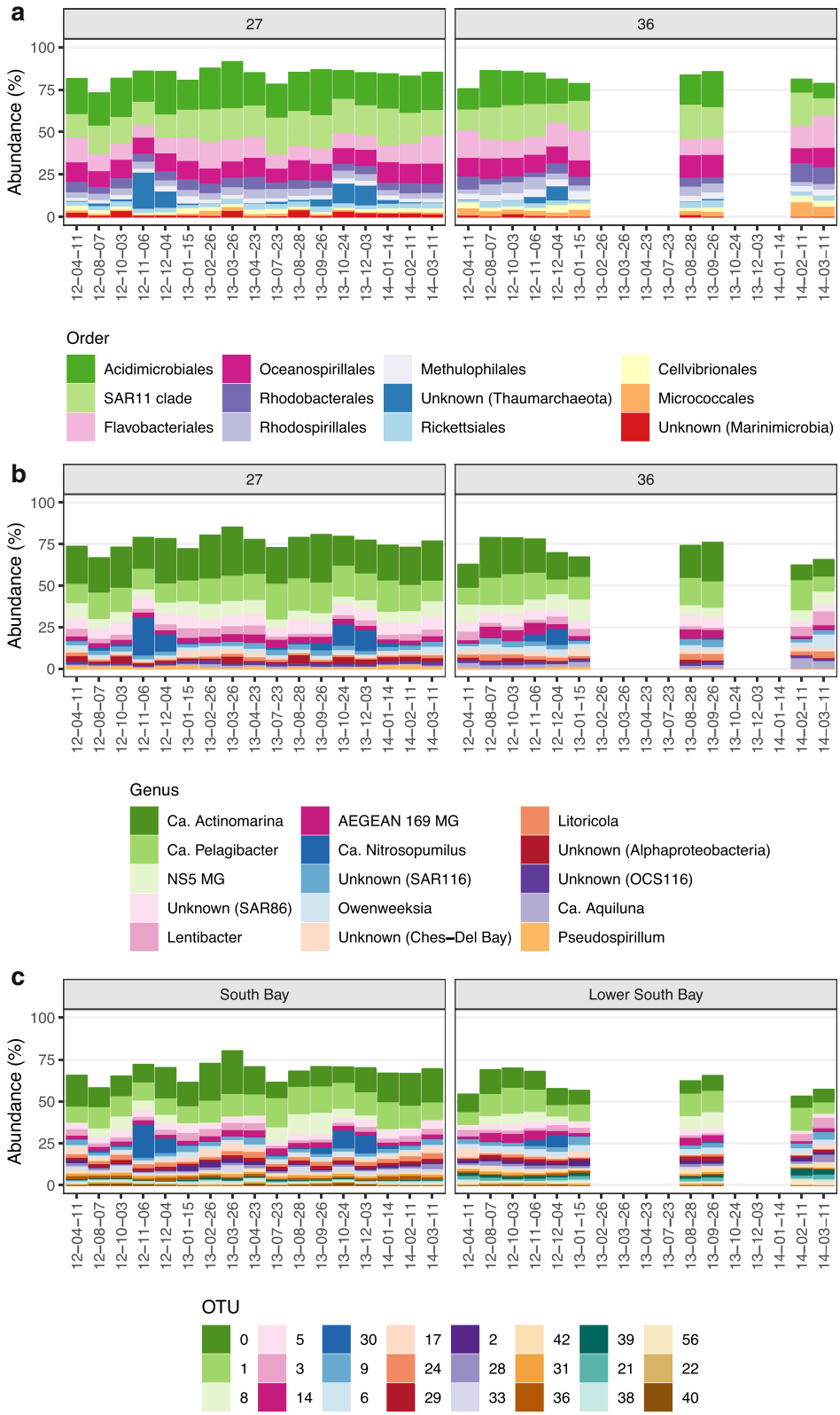

Fig. 7 The % relative abundance of the 12 most abundant orders (**a**), 15 most abundant genera (**b**), and 24 most abundant OTUs (**c**) at 2 stations representative of South Bay (27) and Lower South Bay (36). The *y*-axis shows the relative abundance of taxa in the whole community and the *x*-axis of each panel corresponds to the date (YY-MM-DD) of 20 approximately monthly cruises from April 2012 to March 2014, with the station indicated at the top of the panel. The colors do not correspond between a, b, and c

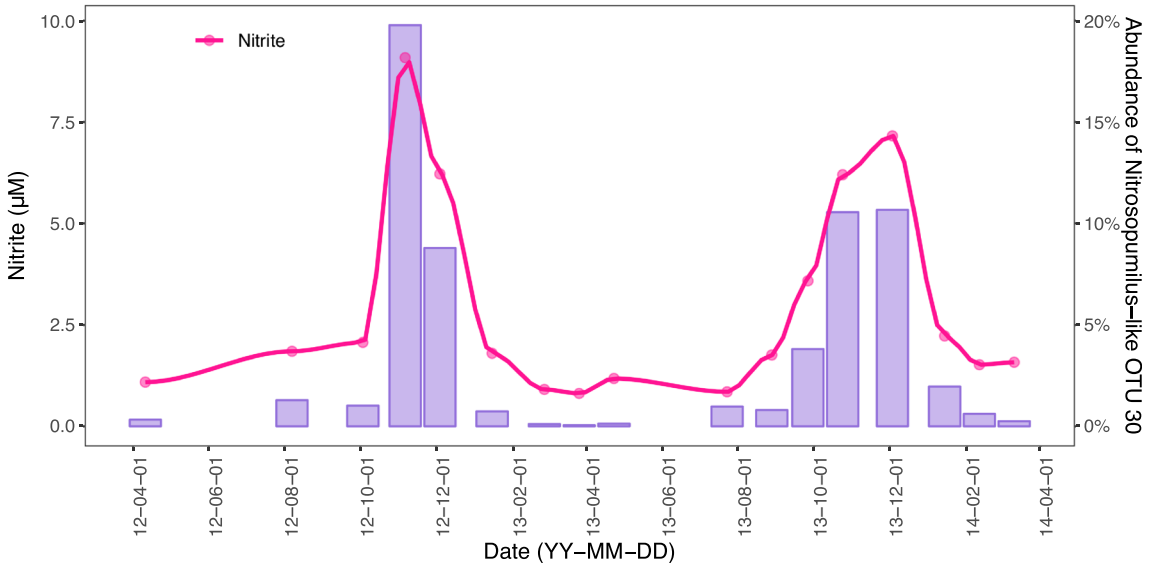

Fig. 8 The % relative abundance of *Nitrosopumilus*-like OTU 30 at South Bay station 27 is shown in purple bars and the corresponding nitrite concentration is shown in pink. The *x*-axis of each panel corresponds to the date (YY-MM-DD) of 20 approximately monthly cruises from April 2012 to March 2014. Note the *y*-axis has two scales, the left corresponding to nitrite concentration and the right to % relative abundance

expected residence time, salinity, temperature, nitrate, and SPM have the highest rank correlation (bioenv, $\rho = 0.59$) and can explain 42.0% of the variation in community dissimilarity in poly/euhaline sites (Supplemental Fig. S8). We did not explicitly measure residence times and rely on values from the literature, highlighting an important need to further explore the impacts of water residence time on brackish community formation in SFB. It has been observed in other estuaries that long residence times do indeed lead to distinct brackish communities [22, 26, 71]. However, we point out that brackish communities in oligo-, meso-, and polyhaline zones are distinct from fresh and marine end-members and also that "species" replacement (rather than "species" loss) is dominant in SFB. Distinct communities are robust in the Bay even though residence times vary from days to months, in agreement with findings that there is a specialized brackish microbiome found in estuaries around the globe [71].

SPM is Important for Both Alpha and Beta Diversity in San Francisco Bay

SPM concentration influences both alpha and beta diversity of bacterio- and archaeoplankton communities throughout SFB. Measures of richness were most strongly correlated with SPM, including the observed OTUs ($r^2 = 0.58$), Chao1 ($r^2 = 0.53$), and Inverse Simpson ($r^2 = 0.14$) (Supplemental Fig. S9). None of these measures were significantly correlated with salinity or library size and were similar whether North and South Bay were analyzed separately or together. In dbRDAs of both North and South Bay, SPM has the second and first largest marginal effect size, respectively (Figs. 2 and 3). SPM could be important for microbial community composition because it provides nutrients for organisms and/or has a long residence time that allows unique, particle-associated communities to form on or around particles in SFB [15]. Indeed, phytoplankton dynamics and bacterioplankton activity have been linked to SPM dynamics in SFB [11, 15, 29, 96] and other estuaries [97–99]. Microbes associated with resuspended sediments could also explain the potential distinction between low- and high-SPM-associated communities. This idea is supported by the negative correlation between observed OTUs and the ratio of active chlorophyll *a* to degraded phaeopigments ($r^2 = 0.35$). The USGS uses a low ratio of active chlorophyll *a* to phaeopigments as an indicator for strong resuspension of bottom sediments, which are rich in degraded algal material. Our study does not differentiate between "free-living" and "particle-associated" organisms as in Hollibaugh et al. [15] and includes all organisms small enough to pass through a 10-µm filter but captured by a 0.22-µm filter. Hollibaugh et al. [15] found free-living and particle-associated communities were very similar in surface waters of SFB, though studies in other estuaries have found differences in community composition between different size fractions [21, 30, 99–101]. While the correlation between SPM and richness in our study could indicate that particle-associated communities (due to either resuspension of sediments or enhanced growth of organisms on/around particles in the water column) are more diverse and distinct from free-living communities at the OTU level, this hypothesis must be explicitly tested. Further studies are necessary in order to directly assess differences between various size fractions of SFB planktonic communities at the OTU level, as well as to

tease apart the relationships between microbial community composition and SPM quantity, quality, and source.

Conclusions

This is the first in-depth spatiotemporal survey of microbial communities in SFB. Overall, we leveraged diversity metrics across and within salinity zones and at different phylogenetic resolution to acquire insights into the biogeography of bacterioplankton and archaeoplankton in North and South SFB. Salinity is a key driver of community structure; however, communities do show beta diversity patterns linked with seasonal variations, regional variations (North vs. South Bay), and changes in SPM concentration. In North Bay, the broader distribution of organisms at coarse taxonomic levels, and the predominance of turnover, low nestedness, and high site specificity of OTUs indicate that closely related organisms have specialized to different salinity ranges along the gradient. Dominant taxonomic groups in SFB (e.g., *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, SAR11) are abundant across the salinity gradient; however, at finer taxonomic levels (e.g., families, genera, and OTUs) within these groups, we see specialization to narrower salinity regimes. Some taxa can be seasonally abundant, including *Synechococcus*, *Ca*. Actinomarina, and especially *Nitrosopumilus*-like organisms, which form an intriguing but previously unrecognized ammonia-oxidizing archaeal "bloom" within this estuary.

Acknowledgments Thank you to Jim Cloern and the Water Quality of San Francisco Bay monitoring group (including but not limited to Tara Schraga, Jessica Dyke, Amy Kleckner, Jennifer Teschler, Charlie Martin, and Jan Thompson) at USGS and the R/V *Polaris* crew for facilitating our participation in numerous cruises. This work was supported in part by NSF CAREER grant OCE-0847266 from the Biological Oceanography program (to CAF), the Stanford McGee research grant (JD), and NSF GRFP and Amherst College Fellowships (to ANR). The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Sequencing thanks to JGI CSP project 503022 to CAF. Special thanks to Tijana Glavino del Rio at JGI.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Nichols FH, Cloern JE, Luoma SN, Peterson DH (1986) The modification of an estuary. Science 231:567–573. https://doi.org/10.1126/science.231.4738.567
2. Kimmerer W (2004) Open water processes of the San Francisco estuary: from physical forcing to biological responses. San Franc Estuary Watershed Sci:2
3. Raimonet M, Cloern JE (2017) Estuary–ocean connectivity: fast physics, slow biology. Glob Change Biol 23:2345–2357. https://doi.org/10.1111/gcb.13546
4. Cloern JE (2001) Our evolving conceptual model of the coastal eutrophication problem. Mar Ecol Prog Ser 210:223–253. https://doi.org/10.3354/meps210223
5. Lucas LV, Koseff JR, Monismith SG, Thompson JK (2009) Shallow water processes govern system-wide phytoplankton bloom dynamics: a modeling study. J Mar Syst 75:70–86. https://doi.org/10.1016/j.jmarsys.2008.07.011
6. Cloern JE, Jassby AD (2012) Drivers of change in estuarine-coastal ecosystems: discoveries from four decades of study in San Francisco Bay. Rev Geophys 50:RG4001. https://doi.org/10.1029/2012RG000397
7. Schraga TS, Cloern JE (2017) Water quality measurements in San Francisco Bay by the U.S. Geological Survey, 1969–2015. Sci Data 4:170098. https://doi.org/10.1038/sdata.2017.98
8. Beck MW, Jabusch TW, Trowbridge PR, Senn DB (2018) Four decades of water quality change in the upper San Francisco Estuary. Estuar Coast Shelf Sci 212:11–22. https://doi.org/10.1016/j.ecss.2018.06.021
9. Cloern JE (2019) Patterns, pace, and processes of water-quality variability in a long-studied estuary. Limnol Oceanogr 64:S192–S208. https://doi.org/10.1002/lno.10958
10. Cloern JE, Jassby AD, Schraga TS, Nejad E, Martin C (2017) Ecosystem variability along the estuarine salinity gradient: examples from long-term study of San Francisco Bay. Limnol Oceanogr 62:S272–S291. https://doi.org/10.1002/lno.10537
11. Cloern JE (1987) Turbidity as a control on phytoplankton biomass and productivity in estuaries. Cont Shelf Res 7:1367–1381. https://doi.org/10.1016/0278-4343(87)90042-2
12. Cloern JE, Dufford R (2005) Phytoplankton community ecology: principles applied in San Francisco Bay. Mar Ecol Prog Ser 285:11–28. https://doi.org/10.3354/meps285011
13. Cloern JE, Jassby AD (2010) Patterns and scales of phytoplankton variability in estuarine–coastal ecosystems. Estuar Coasts 33:230–241. https://doi.org/10.1007/s12237-009-9195-3
14. Sutula M, Kudela R, Hagy JD et al (2017) Novel analyses of long-term data provide a scientific basis for chlorophyll-a thresholds in San Francisco Bay. Estuar Coast Shelf Sci 197:107–118. https://doi.org/10.1016/j.ecss.2017.07.009
15. Hollibaugh JT, Wong PS, Murrell MC (2000) Similarity of particle-associated and free-living bacterial communities in northern San Francisco Bay, California. Aquat Microb Ecol 21:103–114. https://doi.org/10.3354/ame021103
16. Murray AE, Hollibaugh JT, Orrego C (1996) Phylogenetic compositions of bacterioplankton from two California estuaries compared by denaturing gradient gel electrophoresis of 16S rDNA fragments. Appl Environ Microbiol 62:2676–2680
17. Stepanauskas R, Moran MA, Bergamaschi BA, Hollibaugh JT (2003) Covariance of bacterioplankton composition and environmental variables in a temperate delta system. Aquat Microb Ecol 31:14
18. Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. Proc Natl Acad Sci 104:11436–11440. https://doi.org/10.1073/pnas.0611525104
19. Thompson LR, Sanders JG, McDonald D et al (2017) A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551:457–463. https://doi.org/10.1038/nature24621
20. Paver SF, Muratore D, Newton RJ, Coleman ML (2018) Reevaluating the salty divide: phylogenetic specificity of transitions between marine and freshwater systems. mSystems 3:e00232-18. https://doi.org/10.1128/mSystems.00232-18
21. Crump BC, Armbrust EV, Baross JA (1999) Phylogenetic analysis of particle-attached and free-living bacterial communities in the

Columbia River, its estuary, and the adjacent coastal ocean. Appl Environ Microbiol 65:3192–3204
22. Crump BC, Hopkinson CS, Sogin ML, Hobbie JE (2004) Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. Appl Environ Microbiol 70:1494–1505. https://doi.org/10.1128/AEM.70.3.1494-1505.2004
23. Doherty M, Yager PL, Moran MA, Coles VJ, Fortunato CS, Krusche AV, Medeiros PM, Payet JP, Richey JE, Satinsky BM, Sawakuchi HO, Ward ND, Crump BC (2017) Bacterial biogeography across the Amazon River-Ocean Continuum. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.00882
24. Fortunato CS, Crump BC (2011) Bacterioplankton community variation across river to ocean environmental gradients. Microb Ecol 62:374–382. https://doi.org/10.1007/s00248-011-9805-z
25. Fortunato CS, Herfort L, Zuber P, Baptista AM, Crump BC (2012) Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. ISME J 6:554–563. https://doi.org/10.1038/ismej.2011.135
26. Herlemann DP, Labrenz M, Jürgens K et al (2011) Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J 5:1571–1579. https://doi.org/10.1038/ismej.2011.41
27. Mason OU, Canter EJ, Gillies LE, Paisie TK, Roberts BJ (2016) Mississippi river plume enriches microbial diversity in the Northern Gulf of Mexico. Front Microbiol 7. https://doi.org/10.3389/fmicb.2016.01048
28. Liu J, Yu S, Zhao M, He B, Zhang XH (2014) Shifts in archaeaplankton community structure along ecological gradients of Pearl Estuary. FEMS Microbiol Ecol 90:424–435. https://doi.org/10.1111/1574-6941.12404
29. Murrell MC, Hollibaugh JT, Silver MW, Wong PS (1999) Bacterioplankton dynamics in northern San Francisco Bay: Role of particle association and seasonal freshwater flow. Limnol Oceanogr 44:295–308. https://doi.org/10.4319/lo.1999.44.2.0295
30. Satinsky BM, Crump BC, Smith CB, Sharma S, Zielinski BL, Doherty M, Meng J, Sun S, Medeiros PM, Paul JH, Coles VJ, Yager PL, Moran MA (2014) Microspatial gene expression patterns in the Amazon River Plume. Proc Natl Acad Sci 111:11085–11090. https://doi.org/10.1073/pnas.1402782111
31. Apprill A, McNally S, Parsons R, Weber L (2015) Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquat Microb Ecol 75:129–137. https://doi.org/10.3354/ame01753
32. Parada AE, Needham DM, Fuhrman JA (2016) Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol 18:1403–1414. https://doi.org/10.1111/1462-2920.13023
33. Walters RA, Cheng RT, Conomos TJ (1985) Time scales of circulation and mixing processes of San Francisco Bay waters. Hydrobiologia 129:24–36. https://doi.org/10.1007/BF00048685
34. Walters W, Hyde ER, Berg-Lyons D et al (2016) Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. mSystems 1:e00009-15. https://doi.org/10.1128/mSystems.00009-15
35. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. https://doi.org/10.1093/bioinformatics/btq461
36. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066
37. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303
38. Callahan BJ, Sankaran K, Fukuyama JA et al (2016) Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. F1000Research 5:1492. https://doi.org/10.12688/f1000research.8986.1
39. McMurdie PJ, Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. PLOS Comput Biol 10:4f. https://doi.org/10.1371/journal.pcbi.1003531
40. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8
41. McMurdie PJ, Holmes S (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLOS ONE 8:e61217. https://doi.org/10.1371/journal.pone.0061217
42. Oksanen J, Blanchet FG, Friendly M, et al (2018) vegan: Community Ecology Package
43. Battaglia B (1959) Final resolution of the symposium on the classification of brackish waters. Archo Oceanogr Limnol 11(suppl): 243–248
44. Almeida-Neto M, Guimarães P, Guimarães PR et al (2008) A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. Oikos 117:1227–1239. https://doi.org/10.1111/j.0030-1299.2008.16644.x
45. Wear EK, Wilbanks EG, Nelson CE, Carlson CA (2018) Primer selection impacts specific population abundances but not community dynamics in a monthly time-series 16S rRNA gene amplicon analysis of coastal marine bacterioplankton. Environ Microbiol 20:2709–2726. https://doi.org/10.1111/1462-2920.14091
46. Novick E, Senn D (2014) External nutrient loads to San Francisco Bay. San Francisco Estuary Institute, Richmond
47. Hewson I, Fuhrman JA (2004) Richness and diversity of bacterioplankton species along an estuarine gradient in Moreton Bay, Australia. Appl Env Microbiol 70:3425–3433. https://doi.org/10.1128/AEM.70.6.3425-3433.2004
48. Aguirre M, Abad D, Albaina A, Cralle L, Goñi-Urriza MS, Estonba A, Zarraonaindia I (2017) Unraveling the environmental and anthropogenic drivers of bacterial community changes in the Estuary of Bilbao and its tributaries. PLOS ONE 12:e0178755. https://doi.org/10.1371/journal.pone.0178755
49. Baselga A (2010) Partitioning the turnover and nestedness components of beta diversity. Glob Ecol Biogeogr 19:134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x
50. Liu J, Fu B, Yang H, Zhao M, He B, Zhang XH (2015) Phylogenetic shifts of bacterioplankton community composition along the Pearl Estuary: the potential impact of hypoxia and nutrients. Front Microbiol 6. https://doi.org/10.3389/fmicb.2015.00064
51. Mehrshad M, Amoozegar MA, Ghai R, Shahzadeh Fazeli SA, Rodriguez-Valera F (2016) Genome reconstruction from metagenomic data sets reveals novel microbes in the brackish waters of the Caspian Sea. Appl Environ Microbiol 82:1599–1612. https://doi.org/10.1128/AEM.03381-15
52. Hugerth LW, Larsson J, Alneberg J, Lindh MV, Legrand C, Pinhassi J, Andersson AF (2015) Metagenome-assembled genomes uncover a global brackish microbiome. Genome Biol 16:279. https://doi.org/10.1186/s13059-015-0834-7
53. Bouvier TC, del Giorgio PA (2002) Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. Limnol Oceanogr 47:453–470. https://doi.org/10.4319/lo.2002.47.2.0453

54. Zhang Y, Jiao NZ, 焦念志, et al (2006) Contribution of major bacterial groups to bacterial biomass production along a salinity gradient in the South China Sea
55. Kan J, Evans SE, Chen F, Suzuki MT (2008) Novel estuarine bacterioplankton in rRNA operon libraries from the Chesapeake Bay. Aquat Microb Ecol 51:55–66. https://doi.org/10.3354/ame01177
56. Henson MW, Lanclos VC, Faircloth BC, Thrash JC (2018) Cultivation and genomics of the first freshwater SAR11 (LD12) isolate. ISME J 12:1846–1860. https://doi.org/10.1038/s41396-018-0092-2
57. Vergin KL, Beszteri B, Monier A, Cameron Thrash J, Temperton B, Treusch AH, Kilpert F, Worden AZ, Giovannoni SJ (2013) High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. ISME J 7:1322–1332. https://doi.org/10.1038/ismej.2013.32
58. Shaw AK, Halpern AL, Beeson K, Tran B, Venter JC, Martiny JBH (2008) It's all relative: ranking the diversity of aquatic bacterial communities. Environ Microbiol 10:2200–2210. https://doi.org/10.1111/j.1462-2920.2008.01626.x
59. Campbell BJ, Kirchman DL (2013) Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. ISME J 7:210–220. https://doi.org/10.1038/ismej.2012.93
60. Herlemann DPR, Woelk J, Labrenz M, Jürgens K (2014) Diversity and abundance of "Pelagibacterales" (SAR11) in the Baltic Sea salinity gradient. Syst Appl Microbiol 37:601–604. https://doi.org/10.1016/j.syapm.2014.09.002
61. Smith MW, Herfort L, Fortunato CS, Crump BC, Simon HM (2017) Microbial players and processes involved in phytoplankton bloom utilization in the water column of a fast-flowing, river-dominated estuary. MicrobiologyOpen 6:e00467. https://doi.org/10.1002/mbo3.467
62. Williams TJ, Wilkins D, Long E, Evans F, DeMaere MZ, Raftery MJ, Cavicchioli R (2013) The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. Environ Microbiol 15:1302–1317. https://doi.org/10.1111/1462-2920.12017
63. Alonso C, Warnecke F, Amann R, Pernthaler J (2007) High local and global diversity of Flavobacteria in marine plankton. Environ Microbiol 9:1253–1266. https://doi.org/10.1111/j.1462-2920.2007.01244.x
64. Gómez-Pereira PR, Fuchs BM, Alonso C, Oliver MJ, van Beusekom JEE, Amann R (2010) Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. ISME J 4:472–487. https://doi.org/10.1038/ismej.2009.142
65. Yang C, Li Y, Zhou B, Zhou Y, Zheng W, Tian Y, van Nostrand JD, Wu L, He Z, Zhou J, Zheng T (2015) Illumina sequencing-based analysis of free-living bacterial community dynamics during an *Akashiwo sanguine* bloom in Xiamen sea, China. Sci Rep 5:8476. https://doi.org/10.1038/srep08476
66. Bennke CM, Krüger K, Kappelmann L, Huang S, Gobet A, Schüler M, Barbe V, Fuchs BM, Michel G, Teeling H, Amann RI (2016) Polysaccharide utilisation loci of Bacteroidetes from two contrasting open ocean sites in the North Atlantic. Environ Microbiol 18:4456–4470. https://doi.org/10.1111/1462-2920.13429
67. Seo J-H, Kang I, Yang S-J, Cho J-C (2017) Characterization of spatial distribution of the bacterial community in the South Sea of Korea. PLoS ONE 12:e0174159. https://doi.org/10.1371/journal.pone.0174159
68. Ngugi DK, Stingl U (2018) High-quality draft single-cell genome sequence of the NS5 Marine Group from the Coastal Red Sea. Genome Announc 6:e00565–e00518. https://doi.org/10.1128/genomeA.00565-18
69. Glöckner FO, Zaichikov E, Belkova N et al (2000) Comparative 16S rRNA analysis of Lake Bacterioplankton reveals globally distributed phylogenetic clusters including an abundant group of Actinobacteria. Appl Environ Microbiol 66:5053–5065. https://doi.org/10.1128/AEM.66.11.5053-5065.2000
70. Kirchman DL, Dittel AI, Malmstrom RR, Cottrell MT (2005) Biogeography of major bacterial groups in the Delaware Estuary. Limnol Oceanogr 50:1697–1706. https://doi.org/10.4319/lo.2005.50.5.1697
71. Holmfeldt K, Dziallas C, Titelman J, Pohlmann K, Grossart HP, Riemann L (2009) Diversity and abundance of freshwater Actinobacteria along environmental gradients in the brackish northern Baltic Sea. Environ Microbiol 11:2042–2054. https://doi.org/10.1111/j.1462-2920.2009.01925.x
72. Ghai R, Mizuno CM, Picazo A, Camacho A, Rodriguez-Valera F (2014) Key roles for freshwater Actinobacteria revealed by deep metagenomic sequencing. Mol Ecol 23:6073–6090. https://doi.org/10.1111/mec.12985
73. Cloern JE (1996) Phytoplankton bloom dynamics in coastal ecosystems: a review with some general lessons from sustained investigation of San Francisco Bay, California. Rev Geophys 34:127–168. https://doi.org/10.1029/96RG00986
74. Cloern JE (2018) Why large cells dominate estuarine phytoplankton. Limnol Oceanogr 63:S392–S409. https://doi.org/10.1002/lno.10749
75. Li WKW (1998) Annual average abundance of heterotrophic bacteria and Synechococcus in surface ocean waters. Limnol Oceanogr 43:1746–1753. https://doi.org/10.4319/lo.1998.43.7.1746
76. Wang K, Wommack KE, Chen F (2011) Abundance and distribution of Synechococcus spp. and Cyanophages in the Chesapeake Bay. Appl Environ Microbiol 77:7459–7468. https://doi.org/10.1128/AEM.00267-11
77. Ahlgren NA, Rocap G (2012) Diversity and distribution of marine Synechococcus: multiple gene phylogenies for consensus classification and development of qPCR assays for sensitive measurement of clades in the ocean. Aquat Microbiol 3:213. https://doi.org/10.3389/fmicb.2012.00213
78. Xia X, Vidyarathna NK, Palenik B, Lee P, Liu H (2015) Comparison of the seasonal variations of Synechococcus assemblage structures in estuarine waters and coastal waters of Hong Kong. Appl Environ Microbiol 81:7644–7655. https://doi.org/10.1128/AEM.01895-15
79. Ghai R, Mizuno CM, Picazo A et al (2013) Metagenomics uncovers a new group of low GC and ultra-small marine Actinobacteria. Sci Rep 3:srep02471. https://doi.org/10.1038/srep02471
80. Mizuno CM, Rodriguez-Valera F, Ghai R (2015) Genomes of planktonic Acidimicrobiales: widening horizons for marine Actinobacteria by metagenomics. mBio 6:e02083-14. https://doi.org/10.1128/mBio.02083-14
81. Xie W, Luo H, Murugapiran SK, Dodsworth JA, Chen S, Sun Y, Hedlund BP, Wang P, Fang H, Deng M, Zhang CL (2018) Localized high abundance of Marine Group II archaea in the subtropical Pearl River Estuary: implications for their niche adaptation. Environ Microbiol 20:734–754. https://doi.org/10.1111/1462-2920.14004
82. Orellana LH, Ben Francis T, Krüger K, Teeling H, Müller MC, Fuchs BM, Konstantinidis KT, Amann RI (2019) Niche differentiation among annually recurrent coastal Marine Group II Euryarchaeota. ISME J 13:1–13. https://doi.org/10.1038/s41396-019-0491-z
83. Mosier AC, Lund MB, Francis CA (2012) Ecophysiology of an ammonia-oxidizing archaeon adapted to low-salinity habitats. Microb Ecol 64:955–963. https://doi.org/10.1007/s00248-012-0075-1

84. Blainey PC, Mosier AC, Potanina A, Francis CA, Quake SR (2011) Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. PLOS ONE 6:e16626. https://doi.org/10.1371/journal.pone.0016626
85. Santoro AE, Dupont CL, Richter RA, Craig MT, Carini P, McIlvin MR, Yang Y, Orsi WD, Moran DM, Saito MA (2015) Genomic and proteomic characterization of "Candidatus Nitrosopelagicus brevis": An ammonia-oxidizing archaeon from the open ocean. Proc Natl Acad Sci 112:1173–1178. https://doi.org/10.1073/pnas.1416223112
86. Wankel SD, Kendall C, Francis CA, Paytan A (2006) Nitrogen sources and cycling in the San Francisco Bay Estuary: a nitrate dual isotopic composition approach. Limnol Oceanogr 51:1654–1664. https://doi.org/10.4319/lo.2006.51.4.1654
87. Damashek J, Casciotti KL, Francis CA (2016) Variable nitrification rates across environmental gradients in turbid, nutrient-rich estuary waters of San Francisco Bay. Estuaries Coasts 39:1050–1071
88. Simonato F, Gómez-Pereira PR, Fuchs BM, Amann R (2010) Bacterioplankton diversity and community composition in the Southern Lagoon of Venice. Syst Appl Microbiol 33:128–138. https://doi.org/10.1016/j.syapm.2009.12.006
89. Hollibaugh J, Gifford S, Moran MA et al (2014) Seasonal variation in the metratranscriptomes of a Thaumarchaeota population from SE USA coastal waters. ISME J 8:685–698. https://doi.org/10.1038/ismej.2013.171
90. Schaefer SC, Hollibaugh JT (2017) Temperature decouples ammonium and nitrite oxidation in coastal waters. Environ Sci Technol 51:3157–3164. https://doi.org/10.1021/acs.est.6b03483
91. Laperriere SM, Nidzieko NJ, Fox RJ, Fisher AW, Santoro AE (2018) Observations of variable ammonia oxidation and nitrous oxide flux in a eutrophic estuary. Estuar Coasts. 42:33–44. https://doi.org/10.1007/s12237-018-0441-4
92. Hu A, Yang Z, Yu C-P, Jiao N (2013) Dynamics of autotrophic marine planktonic Thaumarchaeota in the East China Sea. PLoS ONE 8:e61087. https://doi.org/10.1371/journal.pone.0061087
93. Kim J-G, Gwak J-H, Jung M-Y, An SU, Hyun JH, Kang S, Rhee SK (2019) Distinct temporal dynamics of planktonic archaeal and bacterial assemblages in the bays of the Yellow Sea. PLOS ONE 14:e0221408. https://doi.org/10.1371/journal.pone.0221408
94. Pitcher A, Wuchter C, Siedenberg K, Schouten S, Sinninghe Damsté JS (2011) Crenarchaeol tracks winter blooms of ammonia-oxidizing Thaumarchaeota in the coastal North Sea. Limnol Oceanogr 56:2308–2318. https://doi.org/10.4319/lo.2011.56.6.2308
95. Wuchter C, Abbas B, Coolen MJL, Herfort L, van Bleijswijk J, Timmers P, Strous M, Teira E, Herndl GJ, Middelburg JJ, Schouten S, Sinninghe Damste JS (2006) Archaeal nitrification in the ocean. Proc Natl Acad Sci U S A 103:12317–12322. https://doi.org/10.1073/pnas.0600756103
96. Hollibaugh JT, Wong PS (1999) Microbial processes in the San Francisco Bay estuarine turbidity maximum. Estuaries 22:848–862. https://doi.org/10.2307/1353066
97. Crump BC, Baross JA, Simenstad CA (1998) Dominance of particle-attached bacteria in the Columbia River estuary, USA. Aquat Microb Ecol 14:7–18. https://doi.org/10.3354/ame014007
98. Smith MW, Zeigler Allen L, Allen AE, Herfort L, Simon HM (2013) Contrasting genomic properties of free-living and particle-attached microbial assemblages within a coastal ecosystem. Aquat Microbiol 4:120. https://doi.org/10.3389/fmicb.2013.00120
99. Wang Y, Pan J, Yang J, Zhou Z, Pan Y, Li M (2019) Patterns and processes of free-living and particle-associated bacterioplankton and archaeaplankton communities in a subtropical river-bay system in South China. Limnol Oceanogr 0: https://doi.org/10.1002/lno.11314
100. Li J-L, Salam N, Wang P-D, Chen LX, Jiao JY, Li X, Xian WD, Han MX, Fang BZ, Mou XZ, Li WJ (2018) Discordance between resident and active bacterioplankton in free-living and particle-associated communities in estuary ecosystem. Microb Ecol 76:637–647. https://doi.org/10.1007/s00248-018-1174-4
101. Dupont CL, Larsson J, Yooseph S, Ininbergs K, Goll J, Asplund-Samuelsson J, McCrow JP, Celepli N, Allen LZ, Ekman M, Lucas AJ, Hagström Å, Thiagarajan M, Brindefalk B, Richter AR, Andersson AF, Tenney A, Lundin D, Tovchigrechko A, Nylander JAA, Brami D, Badger JH, Allen AE, Rusch DB, Hoffman J, Norrby E, Friedman R, Pinhassi J, Venter JC, Bergman B (2014) Functional tradeoffs underpin salinity-driven divergence in microbial community composition. PLOS ONE 9:e89549. https://doi.org/10.1371/journal.pone.0089549

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.