Abundance-Occupancy Relationships Along Taxonomic Ranks Reveal a Consistency of Niche Differentiation in Marine Bacterioplankton With Distinct Lifestyles

Dandan Izabel-Shen[1,2]*, Anna-Lena Höger[2,3] and Klaus Jürgens[2]

[1] *Department of Ecology, Environment and Plant Sciences, Stockholm University, Stockholm, Sweden,* [2] *Department of Biological Oceanography, Leibniz Institute for Baltic Sea Research, Rostock, Germany,* [3] *Department of Applied Biosciences and Process Engineering, Anhalt University of Applied Sciences, Köthen, Germany*

Abundance-occupancy relationships (AORs) are an important determinant of biotic community dynamics and habitat suitability. However, little is known about their role in complex bacterial communities, either within a phylogenetic framework or as a function of niche breadth. Based on data obtained in a field study in the St. Lawrence Estuary, we used 16S rRNA gene sequencing to examine the vertical patterns, strength, and character of AORs for particle-attached and free-living bacterial assemblages. Free-living communities were phylogenetically more diverse than particle-attached communities. The dominant taxa were consistent in terms of their presence/absence but population abundances differed in surface water vs. the cold intermediate layer. Significant, positive AORs characterized all of the surveyed communities across all taxonomic ranks of bacteria, thus demonstrating an ecologically conserved trend for both free-living and particle-attached bacteria. The strength of the AORs was low at the species level but higher at and above the genus level. These results demonstrate that an assessment of the distributions and population densities of finely resolved taxa does not necessarily improve determinations of apparent niche differences in marine bacterioplankton communities at regional scales compared with the information inferred from a broad taxonomic classification.

Keywords: estuary, particle-attached, free-living, 16S amplicon, niche breadth

INTRODUCTION

Aquatic bacterioplankton can either attach to particles or live freely in the water column (Grossart, 2010; Dang and Lovell, 2016). The resulting free-living (FL) and particle-associated (PA) assemblages strongly differ in their community composition and diversity, as shown in several aquatic habitats (e.g., Jackson et al., 2014; Salazar et al., 2015; Milici et al., 2017). For PA assemblages, their community properties depend on the composition of the particles

(Simon et al., 2002) and on the particle types (Simon et al., 2002; Rieck et al., 2015; Bižić-Ionescu et al., 2018), which vary strongly from coastal to open waters. Whereas coastal particles mostly originate from riverine inputs, land runoff, and anthropogenic activities (Rieck et al., 2015; Bižić-Ionescu et al., 2018), oceanic particles are more related to biotic processes such as phytoplankton blooms. In addition, empirical studies suggest that FL and PA bacteria differ in their size and lifestyle and respond differently to environmental fluctuations (Dang and Lovell, 2016, and references therein; Adyari et al., 2020). PA bacteria are generally distinguished by their larger genome sizes and larger number of transporters than found in their FL counterparts (Smith et al., 2013). Among the unique genes PA bacteria are those enabling surface colonization and thereby adaption to environmental fluctuations (DeLong et al., 2006). Generally, the relative contributions of FL and PA bacteria to whole-community functions differ as well. Thus, while FL bacteria dominate in terms of biomass, PA bacteria, despite being less numerous, readily utilize organic aggregates and are thus responsible for a large amount of microbial activity and production (Simon et al., 2002; Dang and Lovell, 2016). Given the distinct roles of FL and PA assemblages in driving different community features, an in-depth understanding of the diversity and dynamics of these ecological groups and of their environmental relationships is a prerequisite for predicting their response to environmental change.

To determine the niche breadth of a species, both the abundance and the occurrence of that species in different environments must be known. Abundance-occupancy relationships (AORs), originally developed for use in macroecology (Gaston et al., 2000; Freckleton et al., 2006), reflect the correlation between the number of sites occupied by a species and the average local abundance of individuals of that species at the occupied site. Positive AORs have been determined for a wide range of organisms (e.g., Gaston et al., 2000; Soininen and Heino, 2005; Barberán et al., 2012; Liu et al., 2015; Shade et al., 2018; Grady et al., 2019). Previous studies highlighted the importance of feedback between local abundance and occupancy in the identification of taxa with strong temporal signatures, thus demonstrating the ability of AORs to reveal large-scale microbial community-environment relationships and taxon persistence (Barberán et al., 2012; Choudoir et al., 2018; Grady et al., 2019). In terms of metacommunity dynamics, changes in local abundance and regional distribution may be a function of a species' niche breadth (Barberán et al., 2012): the broader the niche, the wider the distribution and the greater the local abundance (Lengyel et al., 2020). For microorganisms, even close relatives may be ecologically and physiologically distinct and thereby occupy distinct microbial habitats (Hunt et al., 2008; Koeppel and Wu, 2013; Larkin and Martiny, 2017). Analyzing AORs across taxonomic ranks (from the species to the phylum level) may therefore provide insights into the ecological coherence that underlies habitat-taxon relationships.

The relevance of bacterial taxa as biologically meaningful units is widely discussed in microbial ecology, accompanied by investigations of the ecological consistency of higher taxonomic levels (Koeppel et al., 2008; Philippot et al., 2010;

Koeppel and Wu, 2012). Bacteria assessed at taxonomic ranks above the species level (i.e., genus to phylum) exhibit some ecological coherence, e.g., in their habitat preferences (Philippot et al., 2010) but the phylogenetic depth of that coherence varies among bacterial lineages (Koeppel and Wu, 2012). Moreover, the patterns identified at higher taxonomic ranks may fail to capture the dynamics of the members at lower taxonomic levels (Ruiz-González et al., 2015). Thus, while a broad taxonomic classification may be sufficient to delineate community-environment relationships (Herlemann et al., 2016; Lu et al., 2016), more finely resolved taxonomic analyses may be needed to predict community dynamics (Needham et al., 2017).
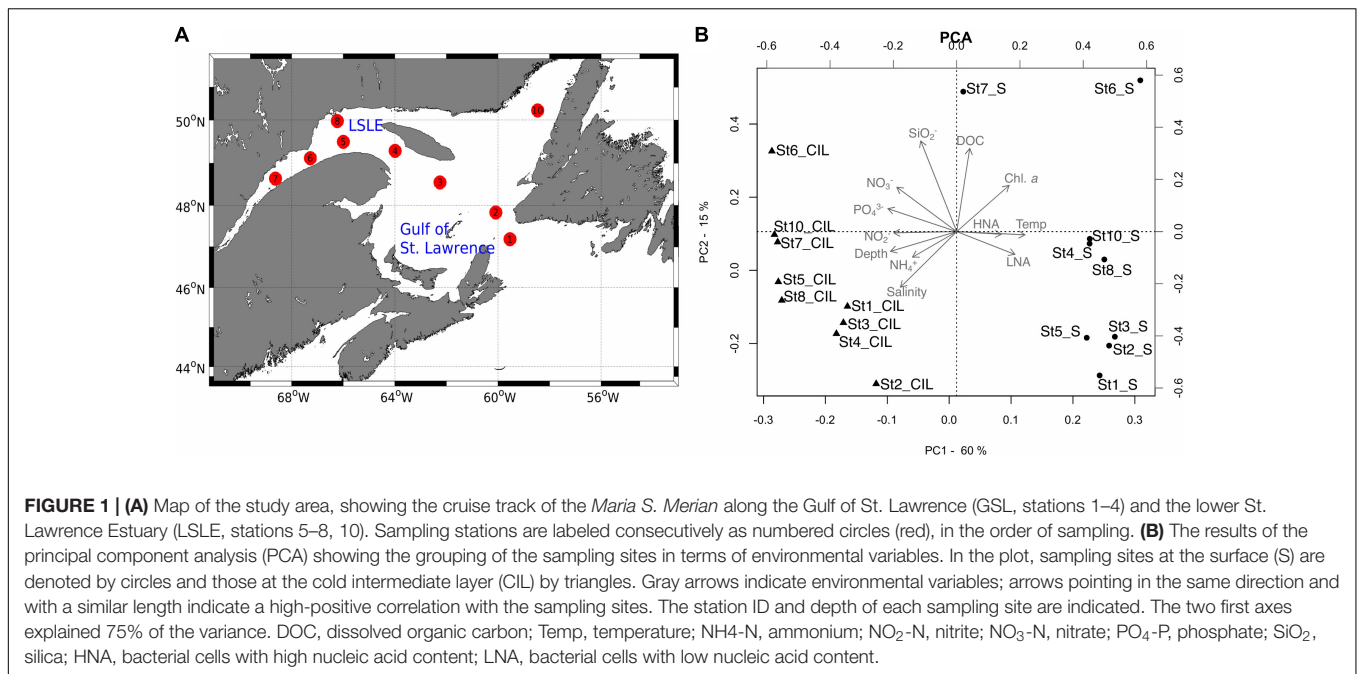
St. Lawrence Estuary (SLE), the deepest and largest estuary in the world, is located on the Canadian east coast and connects the St. Lawrence river over a length of 350 km with the Gulf of St. Lawrence, which separates it from the North Atlantic by nearly 1,000 km (Vincent and Dodson, 1999). The water column of the SLE is stratified in summer, as the deeper cold water is trapped underneath a newly formed warmer layer at the surface (Saucier et al., 2003). Microbially mediated biogeochemical processes and taxon-specific metabolic pathways have been well-documented in the SLE (Bourgoin and Tremblay, 2010; Thibodeau et al., 2010; Ramachandran and Walsh, 2015) but the catalog of whole microbial communities that inhabit the estuary has received less attention. A recent study characterized the PA and FL bacterial communities across the Lower SLE (Cui et al., 2020) but little is known about the alpha-diversity or the relationships between the bacterial community and the environment, and therefore the mechanisms that underlie community assembly.

In this study, the 16S rRNA gene diversity and vertical assembly of FL (0.22–3 µm) and PA (3–200 µm) bacterial assemblages inhabiting the SLE were investigated together with their AORs and the coherence of the observed distributions, both at broad- and fine-scale taxonomic classifications. The following three questions were addressed: (1) Are there vertical patterns (surface and cold intermediate layer) of FL and PA bacterioplankton assembly? If so, what is the relationship between community structure and environmental variables? (2) Among PA and FL communities, can habitat generalists (broad occupancy across sampling sites) and specialists (restricted occupancy to few sites) be distinguished? (3) To what extent can AORs serve as a proxy of the ecological coherence of niche differentiation along taxonomic resolutions?

MATERIALS AND METHODS

Sampling Sites and Water Collection
Water samples were collected during an expedition (MSM 046) of the research vessel Maria S. Merian to the Gulf of St. Lawrence and the lower St. Lawrence Estuary from August 25 to September 4, 2015. The sampling campaign along 2,778 km of the SLE was conducted during a cruise that extended from Halifax to St. John, covering the Lower St. Lawrence Estuary (LSLE) and the Gulf of St. Lawrence (GSL). One of the most distinct features of both regions is a cold intermediate layer (CIL) that develops in summer and is renewed by water from

FIGURE 1 | (A) Map of the study area, showing the cruise track of the *Maria S. Merian* along the Gulf of St. Lawrence (GSL, stations 1–4) and the lower St. Lawrence Estuary (LSLE, stations 5–8, 10). Sampling stations are labeled consecutively as numbered circles (red), in the order of sampling. **(B)** The results of the principal component analysis (PCA) showing the grouping of the sampling sites in terms of environmental variables. In the plot, sampling sites at the surface (S) are denoted by circles and those at the cold intermediate layer (CIL) by triangles. Gray arrows indicate environmental variables; arrows pointing in the same direction and with a similar length indicate a high-positive correlation with the sampling sites. The station ID and depth of each sampling site are indicated. The two first axes explained 75% of the variance. DOC, dissolved organic carbon; Temp, temperature; NH4-N, ammonium; NO₂-N, nitrite; NO₃-N, nitrate; PO₄-P, phosphate; SiO₂, silica; HNA, bacterial cells with high nucleic acid content; LNA, bacterial cells with low nucleic acid content.

different origins (Smith et al., 2006a,b). During the cruise, surface (3 m) and CIL (50–65 m, according to the lowest temperature in the epipelagic zone) samples were collected from nine stations, including determinations of the water chemistry as a measure of environmental heterogeneity (**Figure 1A** and **Supplementary Table 1**).

Sample Collection and Cell Partitioning
For each station, water from the two depths was collected using a CTD rosette sampler and filtered through a 200-μm mesh into 20-L carboys to remove large zooplankton. Conductivity, temperature, and chlorophyll *a* (Chl *a*) fluorescence *in situ* at the corresponding depths were recorded using the CTD rosette during acquisition of the water samples. Fifteen mL of the filtered water of each sample was filtered through GF/F filters (Whatman, Daseel, Germany) and then stored at –20°C until used in the nutrient analysis; another 4 mL was preserved with formaldehyde at a final concentration of 2% for the enumeration of bacterial cells by flow cytometry. These measurements were described in detail in a previous study (Shen et al., 2018b). To separate the PA and FL bacterial communities, cells in 1 L of water were partitioned into two size fractions by sequential filtration of the pre-filtered water through 47-mm diameter Durapore filters with pore sizes of 3.0 μm and 0.22 μm (Millipore, Darmstadt, Germany). The filtration was performed in triplicate for each sample in the same manner. All filters were collected in separate tubes, immediately flash frozen in liquid nitrogen, and stored at –80°C until used for nucleic acid extractions.

DNA Extraction and Illumina Amplicon Sequencing
DNA was extracted from two of the three replicate filters using the Allprep DNA/RNA mini kit (Qiagen, Hilden, Germany)

according to the manufacturer's protocol and then quantified using a NanoDrop 2000 (Thermo Fisher Scientific, Darmstadt, Germany). DNA extraction duplicates were pooled to provide more homogeneous samples, resulting in a total of 36 DNA extracts from distinct water samples. The extracts were PCR-amplified using the bacterial primer pair 341F and 805R (Herlemann et al., 2011). The PCR products were visualized on 1.5% agarose gels before being sent to LGC Genomics GmbH (Berlin, Germany) for high-throughput Illumina sequencing (MiSeq paired-end run with 2 × 300 bp). Each sample was used to obtain duplicate libraries and then sequenced for technical replicates (n = 2).

Sequence Processing and the Reproducibility of Technical Sequencing Replicates
The open-sourced expandable bioinformatics software package MOTHUR v.1.36.1 was used to analyze the sequence data obtained from the Illumina MiSeq platform. The MiSeq Standard Operation Protocol (Kozich et al., 2013) was followed but customized to fit the dataset, as described in Shen et al. (2018b). Briefly, the quality-trimmed remaining sequences were aligned with sequences in the SILVA bacterial reference database v.123 (Quast et al., 2013). The resulting sequences were then checked for chimeras using the UCHIME algorithm (Edgar et al., 2011). All sequences recognized as non-bacterial (chloroplast-mitochondria-Archaea-Eukaryota) or that could not be classified were excluded from the dataset. Finally, the sequences were clustered into operational taxonomic units (OTUs) based on a 99% sequence similarity and using the average neighbor method. All OTUs were classified according to the Bayesian classifier (Wang et al., 2007). The most abundant sequence

in each OTU was chosen as its representative sequence. OTUs that contained only one or two sequences (singletons or doubletons) across all libraries (n = 72, in total) were removed from the dataset. The rationale behind the removal of singleton or doubleton was to reduce the potential sequence errors and improve microbial community assessment but without losing statistical power. In this study, the OTUs were clustered based on a threshold of $\geq$ 99% similarity of the 16S rRNA gene sequences, such that the finely resolved microbial taxa approximately corresponded to bacterial "species" (Kim et al., 2014).

After quality filtering, the 16S rRNA amplicon dataset generated 2, 645, 994 (4, 113, 501 sequences before the omission of single- and doubleton OTUs). Thereafter, 28, 824 OTUs were obtained based on a 99% sequence identity. These OTUs corresponded to 773 genera, 354 families, 177 orders, 94 classes, and 29 phyla. Sequence number varied between the technical replicates for each sample (**Supplementary Table 1**), and such variation might have resulted from the bias of PCR-based library preparation and sequencing errors. Recent work on DNA metabarcoding recommend that technical replicates, e.g., multiple extractions/PCR of the same samples or extracts, are in the experimental workflow (Alberdi et al., 2018; Zinger et al., 2019). Although we did not carry out multiple PCR assays of the same sample during library prep, we did perform an assessment of the reproducibility of the technical replicates (i.e., sequencing replicates, n = 72 in total) for each of the 36 samples. The result showed that despite moderate technical variability, the comparative beta-diversity resembled between the replicates (**Supplementary Figure 1**). We therefore expected that additive merging of sequence reads from the technical replicates would not distort ecological patterns rather achieve a deeper coverage of the respective community. Prior to normalization and downstream analyses, sequence reads from the technical replicates were pooled into one aggregate set of sequences for each sample. Non-normalized technical replicates were pooled into one aggregate set of sequences for each sample. The combined set of sequences was subsampled to 13, 339 sequences (the size of the smallest libraries, see **Supplementary Table 1** for details) to standardize the uneven sequencing effort. In addition to our analyses using a 99% sequence similarity for the OTU classifications, we confirmed that the beta-diversity patterns were maintained, indicating the use of exact sequence variants from the DADA2 pipeline (Callahan et al., 2017) for the analyses. The two beta-diversity patterns (**Supplementary Figure 1** vs. **Supplementary Figure 2**) were found to be statistically indistinguishable based on two separate tests (correlation in a symmetric Procrustes rotation = 0.9711, p-value 0.001 on 999 permutations; Mantel correlation = 0.9565, p-value 0.001 on 999 permutations). Glassman and Martiny (2018) demonstrated that broadscale ecological patterns are robust in terms of the use of exact sequence variants vs. OTUs. Therefore, strong vertical and size-fractionated patterns can be considered robust regardless of whether 99% OTUs or exact sequence variants are selected.

Statistical Analyses

Environmental Relationships Among the Sites and Community Diversity

The variability in the environmental variables (inorganic nutrients, Chl a, dissolved organic carbon (DOC), salinity, temperature, depth, and bacterial abundance) between sampling sites was analyzed by principal component analysis (PCA). Prior to the community analyses, all samples were divided into categorical groups by depth (surface and CIL) and size fraction (FL and PA), resulting in four subcommunities: surface_FL, surface_PA, CIL_FL, and CIL_PA. Within-sample (alpha-) diversity indices, including the observed species richness (S.observed) and evenness (Shannon index/ln S.observed) as well as Faith's phylogenetic diversity (Faith, 1994), were computed from the subsampled dataset. The observed species richness and evenness were calculated using "vegan" package (Oksanen et al., 2020). For phylogenetic diversity, a phylogenetic tree of all OTUs was constructed using FastTree (Price et al., 2009), as implemented in QIIME (Caporaso et al., 2010), and used to calculate the sum of the total phylogenetic branch length of species across samples, using the "picante" package. Student's t-test was used to test the difference in alpha-diversity between the PA and FL communities from the nine stations. Between-sample (beta-) diversity was assessed using the non-metric multidimensional scaling ordination (NMDS) with the Bray-Curtis dissimilarity metric. An analysis of similarity (ANOSIM) was used to test the effects of depth and size fraction on community similarity. A principal coordinates analysis (PCoA) and permutation test were used to analyze the correlation between beta-diversity and environmental factors. Differences in (sub)phylum-level relative abundances between the FL vs. PA communities were determined in a Wilcoxon test.

Inferring Niche Specialization From AORs

The abundance and occupancy patterns of all taxa that made up the subcommunities were analyzed to infer niche specialization patterns in the sampling region. Therefore, two ecological categories with varying degree of niche specialization were defined according to the number of sites at which a taxon was detected (occupancy): taxa present at a minimum of six stations (i.e., an occupancy of > 50%) were identified as habitat generalists, and taxa present at a maximum of two stations (i.e., an occupancy of < 25%) as habitat specialists (adopted from Barberán et al., 2012).

Once the occupancy of a taxon was determined, the mean relative abundance of that taxon was calculated by averaging the aggregated relative abundances across the nine stations. Considering both local factors (abiotic and biotic interactions) and regional factor (dispersal-related processes) are important determinants of abundance and occupancy of a taxon in a locality (Nemergut et al., 2011; Lindström and Langenheder, 2012), we did not exclude the sites at which relative abundance of a taxon was 0% from the calculation of mean relative abundance. Very low mean relative abundances (<0.002%) were excluded from the analysis, as they were assumed to reflect spatial occurrence of taxa by chance (Pandit et al., 2009). This approach was repeatedly

used for all taxonomic ranks, i.e., species (99% OTU sequence similarity), genus, family, order, class, and phylum, for the four subcommunities. In addition, Spearman's rank correlation (*rho*), a non-parametric measure of the degree of association between two variables, was computed to test the correlation between site occupancy and the mean relative abundance. The degree of correlation was interpreted as the strength of the AORs and was used in comparisons across taxonomic ranks. *P*-values indicating significant associations were subjected to Benjamini-Hochberg corrections (Benjamini and Hochberg, 1995) to account for multiple hypothesis testing.

Estimate of Changes in Niche Width Across Taxonomic Ranks

The slope of the AORs, representing the rate of change in abundance vs. occupancy, was used to evaluate the degree of niche width. This analysis was performed for all subcommunities (surface_FL, surface_PA, CIL_FL, and CIL_PA) separately. An analysis of the degree of niche width can offer more information on ecological differences along bacterial taxonomic ranks than provided by phylogeny or taxonomy studies alone (Lu et al., 2019). We therefore asked: (1) whether the degree of niche width would be higher at broad rather than at fine taxonomic resolution, as high taxonomic levels comprise subpopulations with divergent niche preferences, and (2) whether the relationship between niche width and taxonomic rank holds true for communities with different lifestyles and origins. Hence, the values obtained from the slope of the AORs at each taxonomic rank, regardless of lifestyle and depth, were used as replicates in an analysis of variance (ANOVA) to test for significant differences (*P* < 0.05) in the degree of niche width along taxonomic ranks. To fulfill the ANOVA assumption and determine the normality of the data, a Shapiro-Wilk normality test of the residuals in the linear model was conducted. In the case of significant effects of taxonomic rank on the degree of niche width, Tukey's *post hoc* test was used to determine at which level the difference occurred. All statistical analyses and data visualizations were performed in the R environment.

RESULTS

Physiochemical Characteristics and Bacterial Abundances

A suite of contextual data was measured for each sampling site, and the relation between environmental variables and the sampling sites was tested (**Supplementary Table 1** and **Figure 1B**). The PCA revealed a clear separation of the surface and CIL in terms of their water physio-chemistry and prokaryotic cell abundance (**Figure 1B**). The abiotic conditions at the sampling sites were similar at the surface, except at stations 6 and 7, but varied across sites at the CIL. Differences in water temperature, bacterial abundance, and depth had most explanatory values on PC1 (60% variance explained), with contributions from Chl *a*, nitrite, nitrate, and phosphate. This result suggested their important roles in the differentiation of the

two layers. Silica, salinity (and, to a lesser extent, DOC) explained the variation on PC2 (15% variance explained).

Alpha-Diversity

The richness and evenness of the overall communities were greater at the surface than at the CIL (**Figure 2A**). Faith's phylogenetic diversity was significantly higher in the FL than in the PA assemblages across stations (Welch's *t*-tests, *P* < 0.001 for surface; *P* < 0.1 for CIL). At the surface, the species richness of PA was marginally higher than that of FL (*P* < 0.1), while there was no significant difference in the richness of the two groups at the CIL. Conversely, Pielou's evenness was higher for PA than for FL assemblages at the CIL (*P* < 0.05), suggesting the presence of a small number of highly dominant OTUs in the FL groups of the CIL.

Beta-Diversity and Divergent Community Composition

Surface and CIL waters formed diverging environments, as indicated by the PCA; hence, taxonomically distinct community compositions were expected for these habitats. Beta-diversity differed between the two water layers in terms of the taxonomic, weighted resemblance (Bray-Curtis dissimilarity metric, **Figure 2B**). The ANOSIM revealed the significant separation of the four subcommunities from one another by depth (*R* = 0.588, *P* = 0.001) and by size fractionation (*R* = 0.671, *P* = 0.001; **Figure 2B**). Independent of their partitioning as FL or PA, bacterial communities inhabiting the CIL were more tightly clustered than those of the surface, suggesting a greater similarity of CIL communities and less similarity between surface-dwelling communities across stations, especially for PA communities. Salinity, silica levels and temperature significantly correlated with FL assemblages at both surface and CIL (PCoA test, *P* < 0.01 for all), whereas temperature only correlated with surface-dwelling FL (**Table 1**, *P* < 0.01). The levels of inorganic nutrients (NO_3^- and PO_4^{3-}) correlated significantly with the CIL-dwelling FL assemblages (**Table 1**) but the correlations between inorganic nutrients and DOC and PA assemblages were weaker. Cells with low nucleic acid content significantly correlated with surface-dwelling PA assemblages (**Table 1**).

An analysis of the taxonomic composition at the phylum/class level revealed that the relative abundances of the dominant taxa were more variable at surface than at CIL stations (**Figure 3**). Proteobacteria (several subclasses), Bacteroidetes, Actinobacteria, Verrucomicrobia, Cyanobacteria, and Planctomycetes were the dominant bacterial phyla across all stations (**Figure 3**), with Gammaproteobacteria and Deltaproteobacteria predominating at the CIL and Alphaproteobacteria and Actinobacteria in the surface water. Members affiliated with Lentisphaerae, Chloroflexi, and Deferribacteres were mainly detected at the CIL albeit with relatively low abundances (**Figure 3**).

Differences in relative abundance among phyla were determined in comparisons between the size-fractionated categories. For example, the relative abundances of members affiliated with Bacteroidetes, Verrucomicrobia, Planctomycetes,
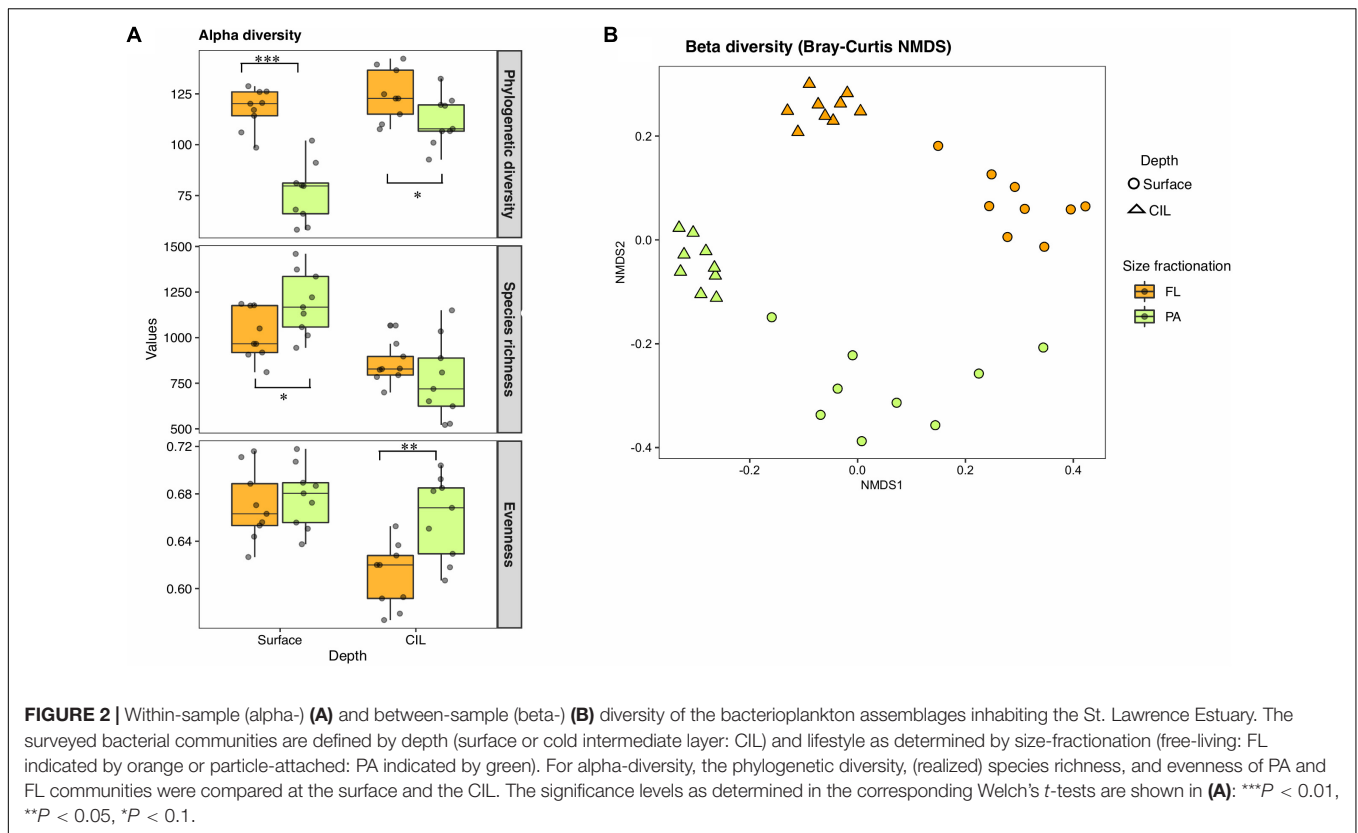
FIGURE 2 | Within-sample (alpha-) **(A)** and between-sample (beta-) **(B)** diversity of the bacterioplankton assemblages inhabiting the St. Lawrence Estuary. The surveyed bacterial communities are defined by depth (surface or cold intermediate layer: CIL) and lifestyle as determined by size-fractionation (free-living: FL indicated by orange or particle-attached: PA indicated by green). For alpha-diversity, the phylogenetic diversity, (realized) species richness, and evenness of PA and FL communities were compared at the surface and the CIL. The significance levels as determined in the corresponding Welch's t-tests are shown in **(A)**: ***$P < 0.01$, **$P < 0.05$, *$P < 0.1$.

TABLE 1 | Principal coordinate analysis (PCoA) and permutation test analyzing the correlation between environmental variables and beta diversity.

Variable tested	Surface_FL		Surface_PA		CIL_FL		CIL_PA	
	R^2	p-value	R^2	p-value	R^2	p-value	R^2	p-value
Salinity	0.8	0.006**	0.191	0.535	0.801	0.001***	0.471	0.104
Temperature	0.82	0.005**	0.583	0.078*	0.275	0.380	0.407	0.173
Chl a	0.2	0.402	0.010	0.975	0.45	0.215	0.450	0.182
NO_3^-	0.55	0.136	0.114	0.712	0.841	0.008**	0.345	0.254
NO_2^-	0.56	0.124	0.157	0.589	0.112	0.746	0.129	0.675
PO_4^{3-}	0.7	0.117	0.323	0.333	0.812	0.007**	0.469	0.162
NH_4^+	0.96	0.116	0.529	0.215	0.306	0.344	0.220	0.507
SiO_2^-	0.89	0.003**	0.254	0.409	0.898	0.001***	0.409	0.217
DOC	0.53	0.093	0.342	0.287	0.035	0.865	0.202	0.434
HNA	0.07	0.866	0.392	0.243	0.002	0.991	0.080	0.813
LNA	0.34	0.272	0.819	0.007**	0.563	0.150	0.091	0.702

Individual communities surveyed are defined by depth (surface or cold intermediate layer: CIL) and size-fractionation (free living: FL or particle-attached: PA).
*Signif. codes: ***$P < 0.001$, **$P < 0.01$, *$P < 0.1$; Number of permutations: 999.*
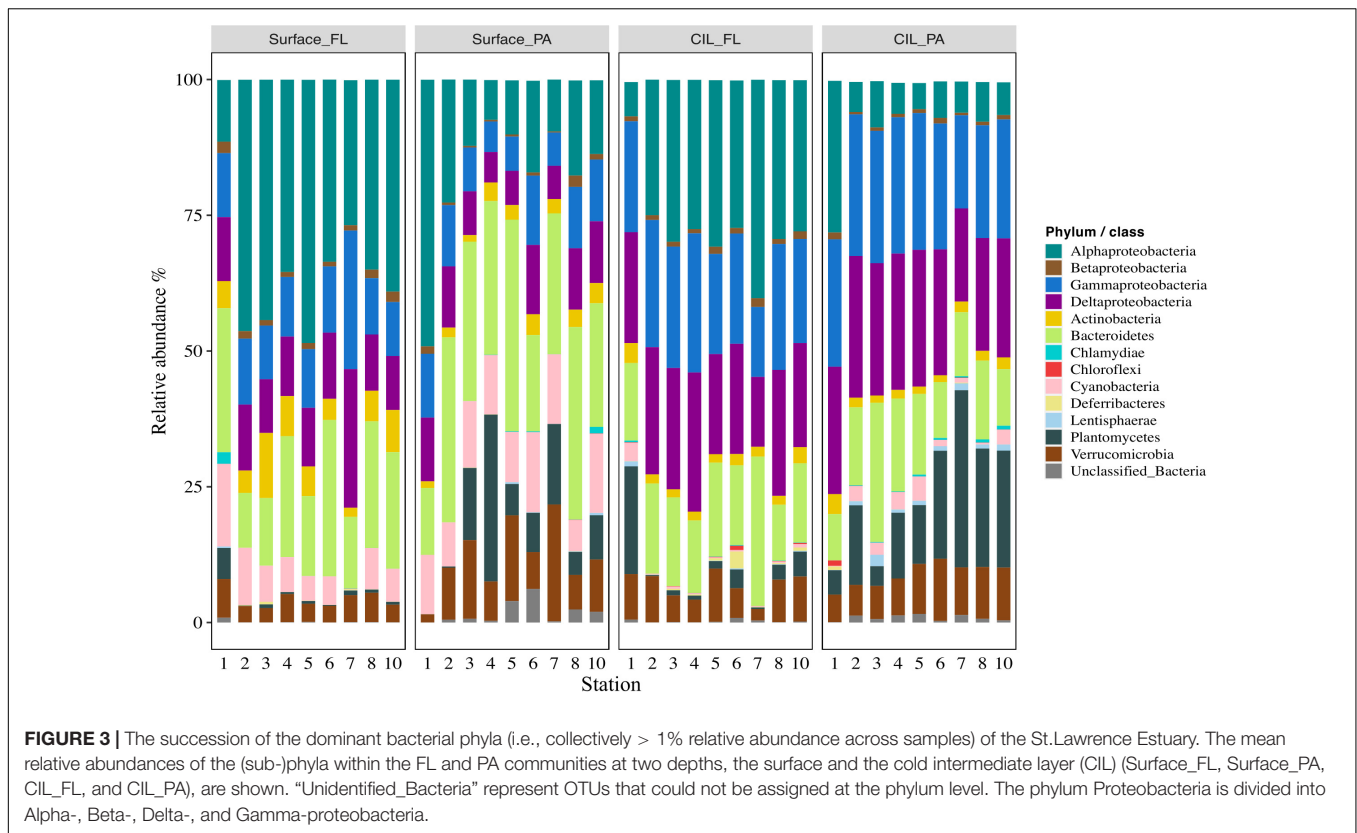*DOC, Dissolved organic carbon; NH4-N, Ammonium; NO_2-N, Nitrite; NO_3-N, Nitrate; PO_4-P, Phosphate; SiO_2, Silica; HNA, bacterial cells with high nucleic acid content; LNA, bacterial cells with low nucleic acid content.*

and Cyanobacteria were significantly higher in PA than in FL communities (for Bacteroidetes and Verrucomicrobia significant only at the surface) (Wilcoxon test, $P < 0.05$; **Supplementary Table 2**). By contrast, the relative abundances of members affiliated with Actinobacteria, Alphaproteobacteria, Betaproteobacteria, and Deltaproteobacteria were significantly higher in FL communities (Wilcoxon test, $P < 0.05$). Despite their overall low contribution, also members affiliated with

Deferribacteres, Firmicutes, and Epsilonproteobacteria differed significantly in relative abundance between PA and FL communities (**Supplementary Table 2**).

AORs Across Taxonomic Ranks and Niche Specialization

In addition to community succession, we explored the AORs of the four subcommunities to obtain insights into
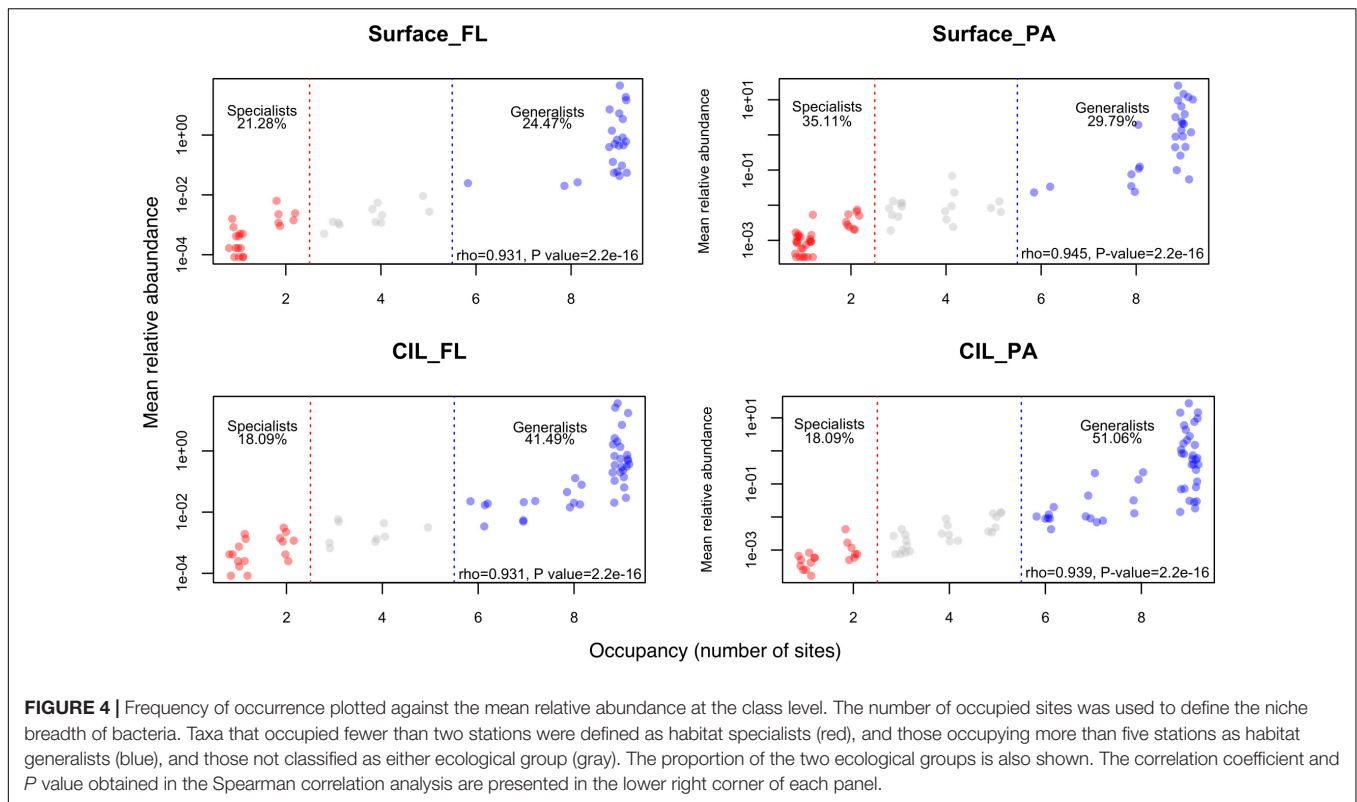
FIGURE 3 | The succession of the dominant bacterial phyla (i.e., collectively > 1% relative abundance across samples) of the St.Lawrence Estuary. The mean relative abundances of the (sub-)phyla within the FL and PA communities at two depths, the surface and the cold intermediate layer (CIL) (Surface_FL, Surface_PA, CIL_FL, and CIL_PA), are shown. "Unidentified_Bacteria" represent OTUs that could not be assigned at the phylum level. The phylum Proteobacteria is divided into Alpha-, Beta-, Delta-, and Gamma-proteobacteria.

niche specialization patterns across taxonomic ranks. The results showed a clear trend in the proportions of the two ecological categories along the taxonomic scale (**Figure 4**: class level; **Supplementary Figure 3**: other taxonomic levels and **Supplementary Table 3**). Niche specialization varies as a function of taxonomic resolution. The niche specialization of individuals when considered at the lower taxonomic levels (i.e., species, genus, order levels), should therefore be finer than that of class or phylum to which they belong. For example, in the case of Surface_Fl subcommunities, three bacterial orders (*Corynebacteriales*, *Frankiales*, *Propionibacteriales* affiliated with Actinobacterial class) occupied seven sites, respectively (**Supplementary Figure 3**, order level; yellow highlight in **Supplementary Table 4**); however, they had occupancy at 9 stations when their occupancy was scaled to the class level Actinobacteria (the case of Surface_FL, **Figure 4**). The number of generalists increased with increasing phylogenetic distance, whereas the opposite was determined for specialists. Additionally, the increase in generalists was larger for PA than for FL communities. This pattern was inverted for specialists, with a more pronounced decrease for FL than for PA communities (**Supplementary Table 3**). Regardless of the taxonomic rank, the AORs were positive, which suggested that species with low abundances tended to have narrower occupancies and those with high abundances wider occupancies (**Figure 4** and **Supplementary Table 5**). Moreover, the taxonomic composition of specialists vs. generalists differed profoundly for PA and FL bacteria (**Supplementary Figure 4**). The compositional changes

of generalists at the class level generally resembled those at the phylum level across depths.

The strength of the AORs of the PA and FL communities across taxonomic ranks was also considered. We hypothesized that the apparent niche differences of natural bacterial communities would be more clearly defined by the distributions and population densities of finely resolved taxa than by inferences based on a broadly resolved taxonomic classification. Thus, AOR strength was expected to decrease when a broad taxonomic classification was used. However, while the AORs of the surveyed bacterial communities were positive across all taxonomic ranks (**Figure 5**, Spearman coefficient > 0), contrary to our expectation, AOR strength was lowest at the species level and increased at broader taxonomic levels (Spearman coefficient: 0.626–0.719), remaining more or less constant from the genus to the phylum level (**Figure 5** and **Supplementary Table 6**).

Degree of Niche Width Inferred From AORs

Both the slope of the AORs and the taxonomic resolution were analyzed to determine whether the degree of niche width of the four subcommunities increased with taxonomic rank. The degree of niche width of the subcommunities increased exponentially with broadening taxonomic resolution (**Figure 6**, ANOVA: $P = 2.42 \times 10^{-6}$); thus, increases in the phylogenetic distance resulted in exponential increases in the niche width of the communities. The deviation between communities began at

FIGURE 4 | Frequency of occurrence plotted against the mean relative abundance at the class level. The number of occupied sites was used to define the niche breadth of bacteria. Taxa that occupied fewer than two stations were defined as habitat specialists (red), and those occupying more than five stations as habitat generalists (blue), and those not classified as either ecological group (gray). The proportion of the two ecological groups is also shown. The correlation coefficient and *P* value obtained in the Spearman correlation analysis are presented in the lower right corner of each panel.

the genus level and became progressively larger until the phylum level, with a significantly large difference at the phylum or class level (**Figure 6**, inset; Tukey's *post hoc*, $P < 0.001$). The increase was quantified by calculating the fold change between every two consecutive taxonomic ranks (**Supplementary Table 7B**), which again showed that the largest change in niche width occurred between the class and phylum levels. This pattern repeated across four subcommunities with a similar magnitude. Furthermore, the niche index of FL communities deviated from that of the PA communities from the species to the order level (**Figure 6**).
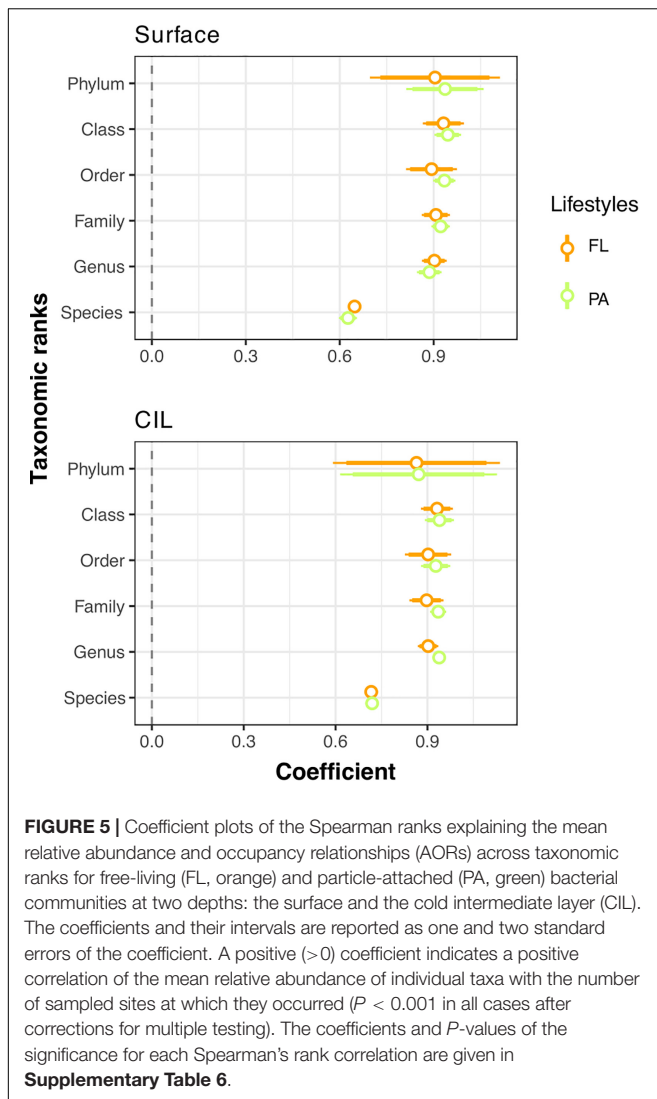
DISCUSSION

Vertical Shifts in Community Characteristics

A strong stratification of the SLE, with pronounced abiotic differences between the surface and the CIL, was determined during our sampling campaign, consistent with previous reports (Vincent and Dodson, 1999; Elliott and McLusky, 2002). Inorganic nutrient concentrations were higher at the sampling sites of the LSLE than at those located in the GSL, attributable to the nutrient-rich water upwelled and adverted from upstream regions (Therriault and Levasseur, 1985) and to the frequent tidal events in the LSLE (Winkler et al., 2003; Dinauer and Mucci, 2017). Bacterial species richness and evenness were higher at the surface than in the CIL (**Figure 2A**). The low level of community diversity at the CIL was likely due to the near-freezing water temperature, given that low temperatures suppress the activity

and richness of bacterial communities (Pomeroy and Wiebe, 2001; Kellogg and Deming, 2009).

Community composition was more similar among CIL stations than among surface stations. On the horizontal scale, surface stations were more divergent from each other than those within CIL in terms of their association with abiotic and biotic parameters (**Figure 1B**), suggesting that the conditions at the surface were more heterogenous across stations than that in the CIL. The heterogeneity in surface water of the LSLE and Gulf of SLE was most likely due to different stages of phytoplankton blooms and thereby different levels of primary production in summer, as previously reported (Zakardjian et al., 2000; Archambault et al., 2010). In accordance with this, we observed a higher variability of e.g., Chl *a* (0.48–2.43 μM/mL), DOC (5.1–19.60 μM/mL), high-nucleic acid (HNA) cells (1.17–8.75 × 10^6 cells/mL) at the surface stations than CIL (0.06–1.15 μM/mL, 4.70–16.20 μM/mL, 1.64–2.59 × 10^5 cells/mL, respectively, **Supplementary Table 1**). Although we did not capture dynamics of all the phytoplankton in this study, cyanobacterial population dominated 10–22% of individual assemblages in surface but decreased to 2% in the CIL (**Figure 3**). As such, DOC released from phytoplankton by direct excretion or through trophic interactions sustains the growth of diverse bacterial assemblages in the surface water (Azam et al., 1983). On a vertical scale, dispersion between the surface and CIL is limited, including vertical mixing not only of the respective environmental matrices but also of microbial cells from each one. It was therefore unlikely that bacteria inhabiting the surface water were able to colonize the CIL.

FIGURE 5 | Coefficient plots of the Spearman ranks explaining the mean relative abundance and occupancy relationships (AORs) across taxonomic ranks for free-living (FL, orange) and particle-attached (PA, green) bacterial communities at two depths: the surface and the cold intermediate layer (CIL). The coefficients and their intervals are reported as one and two standard errors of the coefficient. A positive (>0) coefficient indicates a positive correlation of the mean relative abundance of individual taxa with the number of sampled sites at which they occurred ($P < 0.001$ in all cases after corrections for multiple testing). The coefficients and P-values of the significance for each Spearman's rank correlation are given in **Supplementary Table 6**.

Distinct Community Composition of FL and PA Bacteria

The higher phylogenetic diversity of FL than PA communities suggested that the former harbor more phylogenetically diverse groups of bacteria (**Figure 2A**). However, the species richness and evenness of PA were greater than that of FL at the surface and CIL, respectively. These results are supported by some (Crespo et al., 2013; Mohit et al., 2014; Rieck et al., 2015) but not all studies, with most of the latter reporting a higher diversity for FL bacteria (Acinas, 1999; Ghiglione et al., 2009; Kellogg and Deming, 2009). The inconsistence across studies can be explained by the environmentally (locality-) dependent origin and material composition of organic particles (Simon et al., 2002; Mohit et al., 2014; Bižić-Ionescu et al., 2018). The strong effects of the measured environmental variables on FL but not on PA assemblages (**Table 1**) indicated differences in the environmental interactions of species from different lifestyle groups. Thus, PA bacteria may have been

influenced by environmental factors that were not assessed in this study, such as the concentrations of particulate organic matter (Mohit et al., 2014). Estuarine systems are characterized by large inflows of terrestrial materials and allochthonous organic matter (Dinauer and Mucci, 2017), which may have contributed to the high bacterial richness of the PA fraction in the SLE.

In accordance with earlier findings in marine and estuarine systems (Crump et al., 1999; Salazar et al., 2015; Mestre et al., 2017; Cui et al., 2020), the composition of FL bacteria was both taxonomically and phylogenetically distinct from that of PA bacteria (**Figure 2B** and **Supplementary Figure 1**, Unifrac). Bacteroidetes, Planctomycetes, and Verrucomicrobia accounted for a larger proportion of the PA fraction than of the FL fraction. These groups have been described previously as PA indicators (Salazar et al., 2015). For example, a strong particle affinity of members of Planctomycetes was reported for biofilm-forming communities on microplastic particles (Wiegand et al., 2020), in the low-saline communities of the Baltic Sea (Rieck et al., 2015), and in deep-sea prokaryotic assemblages (Salazar et al., 2015). In our study, Cyanobacteria were more abundant in the PA than in the FL fractions at all stations, consistent with previous reports of abundant cyanobacterial populations in the LSLE (Cui et al., 2020) but also in the GSL, where the salinity is higher. The presence of cyanobacterial populations on (sinking) particles was also described (Farnelid et al., 2019; Yang et al., 2019). In our study, the high abundance of PA cyanobacteria in the surface layer and in the CIL points to potential bloom conditions in the SLE and thus the potential export of organic materials from the surface to deeper-water habitats via vertical dispersion in the form of particle attachment. Conversely, the high abundances of Alphaproteobacteria and Actinobacteria in the FL fraction in surface water (**Figure 3**) were consistent with the preferentially pelagic and FL lifestyle of these taxa (Dang and Lovell, 2016).

AORs and Habitat Specialization

By applying an established macroecological approach (Gaston et al., 2000) to consider both the occupancy and abundance patterns of bacterial assemblages inhabiting the SLE, we were able to discern two ecological groups: (1) habitat generalists, which tended to be regionally abundant, and (2) habitat specialists, which were regionally rare in all cases, regardless of the taxonomic depth and size fractionation of the bacterioplankton (**Figure 4** and **Supplementary Figure 3**). However, data obtained during field and experimental studies demonstrated that, in bacterial communities, habitat specialists were more locally abundant than generalists (Barberán et al., 2012; Logares et al., 2013; Shen et al., 2018a). The discrepancy can be explained by difference in habitat differentiation, as SLE habitats were less well-differentiated than those surveyed in previous studies of environmental gradients, such that habitat filtering in the SLE was not strong enough to promote the growth of specialists with a high growth rate. These results suggest that definitions of specialists and generalists based on habitat occupancy will depend on the environmental conditions of the

FIGURE 6 | Niche width index (derived from the slope of the logarithm of the abundance values plotted against the number of occupied sites) of the subcommunities along hierarchical taxonomic ranks and the corresponding median rate at each taxonomic rank (inset), without evolutionary information on the bacteria. The different colors indicate the different lifestyles, and the different symbols indicate the different water depths. An analysis of variance (ANOVA) was used to test for significant differences ($P < 0.05$) in the niche width index as a function of the taxonomic level. The subcommunities function as replicates for the ANOVA; thus, the values obtained from the subcommunities were used to calculate the variance of these replicates. Tukey's *post hoc* test was used to determine the taxonomic level at which the differences occurred. The corresponding homogeneous groups are indicated by a, b and c. The values of the slopes of the AORs of the individual communities are given in **Supplementary Table 7A**.

sampling area. It has also been suggested that habitat specialists tend to outnumber generalists in ecological communities (Guo et al., 2000; Van der Gast et al., 2011). However, at least in our study, this tendency skewed toward high taxonomic ranks. Among communities above the order level, generalists made up a larger proportion than specialists (**Supplementary Figures 3E,F**), with the most pronounced difference occurring at the phylum level. Higher taxonomic ranks are made up of subgroups characterized by increased phylogenetic distances and therefore possibly divergent ecological preferences (Philippot et al., 2010). Overall, our results demonstrate that the number and taxonomic resolution of habitat generalists vs. specialists should be taken into account in studies of community dynamics and habitat suitability.

In addition, a significant positive relationship was determined between the mean relative abundance and occupancy for the subcommunities surveyed at all taxonomic levels (**Supplementary Figure 4**). Similar relationships were identified both in lake (Liu et al., 2015) and gut microbial communities in the zebrafish (Burns et al., 2016). The AORs were stronger at the class level and weaker at the species level (OTUs at 99% sequence similarity). The high probability of randomness at finer taxonomic scales, such as the species level, may obscure community-environment relationships (Lu et al., 2016).

Ecological Coherence of Niche Width in FL and PA Assemblages

Our data revealed an exponential increase in niche width index (a proxy of ecological niche) with increasing phylogenetic distance, with the sharpest increase occurring between class and phylum

(**Figure 6** and **Supplementary Table 7B**). This indicates that bacterial taxa possess an ecological coherence that weakens at higher taxonomic ranks. Philippot et al. (2010) came to a similar conclusion in their investigation of the habitat/phylotype association of a set of microbial strains whose genomes had been sequenced. In contrast to our study, the authors used an approach focused on habitat similarity rather than the divergences of habitats and niche width. Ultimately, they found a negative correlation between habitat similarity and taxonomic ranks and thus reached the same conclusion as resulting from our study, i.e., that the ecological consistency of this relation diminishes above the class level. At higher taxonomic ranks (i.e., class or even phylum level), the niche indices of the subcommunities varied as a function of water depth rather than lifestyle, observed at finer taxonomic scales. These findings suggest that niche differentiation is a phylogenetically conserved tendency of both PA and FL bacteria, as noted elsewhere (Mestre et al., 2017), and that the effect of water depth on the entire communities within each lifestyle is stronger when broad taxonomic scales are applied.

Caveats and Recommendations for Future Work

Our study had several limitations. First, a 99% sequence similarity cutoff was used to aggregate OTUs at the species level, although a cutoff of 97% is more common. This might have led to an overestimation of taxa, given the inherent sequencing errors and biases. Hence, the OTU-based classification should be interpreted as describing closely related bacterial lineages

rather than species. Second, FL and PA bacteria were partitioned based on two cell size categories (0.2–3 μm and > 3 μm). Since there is no consensus on how to most effectively separate these groups, caution is needed in comparing the community patterns identified in our study vs. in previous studies, in which different size-fractionation methods were used to partition FL and PA bacteria. Nevertheless, the methods chosen for this study are commonly used in microbial ecology and they provided insights into community-environment relationships in the SLE and the vertical assembly of bacterial communities that differ in their lifestyles.

CONCLUSION

Our study showed that the FL fraction consisted of phylogenetically more diverse bacterial populations than the PA fraction, but vertical patterns were detected within each fraction. The significant and positive AORs identified across all taxonomic ranks of bacteria demonstrated an ecologically conserved trend for both lifestyles, even at the small regional scale of the study. AOR strength was low at the species level but increased steadily with progressively higher taxonomic rank, irrespective of bacterial lifestyle. Thus, in studies of bacterial assemblages at regional scales, analyses of high taxonomic ranks may already provide a reasonable understanding of contemporary patterns of their abundance and occurrence. Finally, our analysis of the degree of niche width inferred from the AOR slopes along taxonomic ranks revealed the consistent niche differentiation of bacterioplankton with different lifestyles, with greater differentiations at higher (class and phylum) taxonomic levels. Collectively, our results corroborate previous work (e.g., Shade et al., 2018) in that they extend the principles developed for plants and animals to an understanding of the dynamics and distribution of complex natural microbial communities. Thus, not only is a FL or PA lifestyle a phylogenetically conserved trait for pelagic marine bacterioplankton (Salazar et al., 2015), but the niche differences and distributions of these two groups at regional scales are maintained over broad taxonomic resolution. Future work should ascertain whether these patterns also describe other microbial communities classified by multiple size-fractionations and whether the relationship is scale-dependent.

DATA AVAILABILITY STATEMENT

The FASTQ files and associated metadata are publicly available at the European Nucleotide Archive under the accession number PRJEB30352. The R scripts and computing notes for this study are available on Zenodo (https://doi.org/10.5281/zenodo.4743168) with supplement to Github (https://github.com/IzabelShen/Abundance-occupancy).

AUTHOR CONTRIBUTIONS

DI-S and KJ conceived the study. DI-S conducted field and laboratory work, data analysis, and wrote a first version of the manuscript. A-LH processed samples and performed data analysis. All authors discussed the results and commented on the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.690712/full#supplementary-material

REFERENCES

Acinas, S. G. (1999). Diversity of free-living and attached bacteria in offshore Western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. *Appl. Environ. Microbiol.* 65, 514–522. doi: 10.1128/aem.65.2.514-522.1999

Adyari, B., Shen, D., Li, S., Zhang, L., Rashid, A., Sun, Q., et al. (2020). Strong impact of micropollutants on prokaryotic communities at the horizontal but not vertical scales in a subtropical reservoir, China. *Sci. Total Environ.* 721:137767. doi: 10.1016/j.scitotenv.2020.137767

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., and Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods Ecol. Evol.* 9, 134–147. doi: 10.1111/2041-210x.12849

Archambault, P., Snelgrove, P. V. R., Fisher, J. A. D., Gagnon, J., Garbary, D. J., Harvey, M., et al. (2010). From sea to sea: Canada's three oceans of biodiversity. *PLoS One* 5:e12182. doi: 10.1371/journal.pone.0012182

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A., and Thingstad, F. (1983). The ecological role of water-column microbes in the Sea. *Mar. Ecol. Prog. Ser.* 10, 257–263. doi: 10.3354/meps010257

Barberán, A., Bates, S. T., Casamayor, E. O., and Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 6, 343–351. doi: 10.1038/ismej.2011.119

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bižić-Ionescu, M., Ionescu, D., and Grossart, H. (2018). Organic particles: heterogenous hubs for microbial interactions in aquatic ecosystems. *Front. Microbiol.* 9:2569. doi: 10.3389/fmicb.2018.02569

Bourgoin, L., and Tremblay, L. (2010). Bacterial reworking of terrigenous and marine organic matter in estuarine water columns and sediments. *Geochim. Cosmochim. Acta* 74, 5593–5609. doi: 10.1016/j.gca.2010.06.037

Burns, A. R., Stephens, W. Z., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., et al. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J.* 10, 655–664. doi: 10.1038/ismej.2015.142

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639–2643. doi: 10.1038/ismej.2017.119

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throuhput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Choudoir, M. J., Barberán, A., Menninger, H. L., Dunn, R. R., and Fierer, N. (2018). Variation in range size and dispersal capabilities of microbial taxa. *Ecology* 99, 322–334. doi: 10.1002/ecy.2094

Crespo, B. G., Pommier, T., Fernandez-Gomez, B., and Pedros-Alio, C. (2013). Taxonomic composition of the particle-attached and free-living bacterial assemblages in the Northwest Mediterranean Sea analyzed by pyrosequencing of the 16S rRNA. *Microbiologyopen* 2, 541–552. doi: 10.1002/mbo3.92

Crump, B. C., Armbrust, E. V., and Baross, J. A. (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.* 64, 3192–3204. doi: 10.1128/aem.65.7.3192-3204.1999

Cui, T., Dawson, T. J., McLatchie, S., Dunn, K., Bielawski, J., and Walsh, D. A. (2020). Modelling free-living and particle-associated bacterial assemblages across the deep and hypoxic Lower St. Lawrence Estuary. *mSphere* 5:e00364–20.

Dang, H., and Lovell, C. (2016). Microbial surface colonization and biofilm development in marine environments. *Microbiol. Mol. Biol. Rev.* 80, 91–138. doi: 10.1128/mmbr.00037-15

DeLong, E. F., Preston, C. M., Mincer, T., Rich, V., Hallam, S. J., Frigaard, N. U., et al. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311, 496–503. doi: 10.1126/science.1120250

Dinauer, A., and Mucci, A. (2017). Spatial variability in surface-water pCO2 and gas exchange in the world's largest semi-enclosed estuarine system: St. Lawrence Estuary (Canada). *Biogeosciences* 14, 3221–3237. doi: 10.5194/bg-14-3221-2017

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Elliott, M., and McLusky, D. S. (2002). The need for definitions in understanding estuaries. *Estuar. Coast. Shelf Sci.* 55, 815–827. doi: 10.1006/ecss.2002.1031

Faith, D. P. (1994). "Phylogenetic diversity: a general framework for the prediction of feature diversity," in *Systematics and Conservation Evaluation*, eds P. L. Forey, C. J. Humphries, and R. I. Vane-Wright (Oxford: Clarendon Press), 241–268.

Farnelid, H., Turk-Kubo, K., Ploug, H., Ossolinski, J. E., Collins, J. R., Van Mooy, B. A. S., et al. (2019). Diverse diazotrophs are present on sinking particles in the North Pacific Subtropical Gyre. *ISME J.* 13, 170–182. doi: 10.1038/s41396-018-0259-x

Freckleton, R. P., Noble, D., and Webb, T. J. (2006). Distributions of habitat suitability and the abundance-occupancy relationship. *Am. Nat.* 167, 260–275. doi: 10.2307/3491266

Gaston, K. J., Blackburn, T. M., Greenwood, J. J. D., Gregory, R. D., Quinn, R. M., and Lawton, J. H. (2000). Abundance-occupancy relationships. *J. Appl. Ecol.* 37, 39–59.

Ghiglione, J. F., Conan, P., and Pujo-Pay, M. (2009). Diversity of total and active free-living vs. particle-attached bacteria in the eutrophic zone of the NW Mediterranean Sea. *FEMS Microbiol. Ecol.* 299, 9–21. doi: 10.1111/j.1574-6968.2009.01694.x

Glassman, S. I., and Martiny, J. B. H. (2018). Broadscale ecological patterns are robust to use of exact sequence variants versus operational taxonomic units. *mSphere* 3:e00148–18.

Grady, K. L., Sorensen, J. W., Stopnisek, N., Guittar, J., and Shade, A. (2019). Assembly and seasonality of core phyllosphere microbiota and perennial biofuel crops. *Nat. Commun.* 10:4135.

Grossart, H. (2010). Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. *Environ. Microbiol.* 2, 706–714. doi: 10.1111/j.1758-2229.2010.00179.x

Guo, Q., Brown, J. H., and Valone, T. J. (2000). Abundance and distribution of desert annuals: are spatial and temporal patterns related? *Ecology* 88, 551–560. doi: 10.1046/j.1365-2745.2000.00466.x

Herlemann, D. P. R., Lundin, D., Andersson, A. F., Labrenz, M., and Jürgens, K. (2016). Phylogenetic signals of salinity and season in bacterial community composition across the salinity gradient of the Baltic Sea. *Front. Microbiol.* 7:1883. doi: 10.3389/fmicb.2016.01883

Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., Andersson, A. F., et al. (2011). Transition in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* 5, 1571–1579. doi: 10.1038/ismej.2011.41

Hunt, D. E., David, L. A., Gevers, D., Preheim, S. P., Alm, E. J., and Polz, M. F. (2008). Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320, 1081–1085. doi: 10.1126/science.1157890

Izabel-Shen, D., Höger, A., and Jürgens, K. (2021). Abundance-occupancy relationships along taxonomic ranks reveal a consistency of niche differentiation in marine bacterioplankton with distinct lifestyles. *bioRxiv* [Preprint]. doi: 10.1101/2021.04.09.439152

Jackson, C. R., Millar, J. J., Payne, J. T., and Ochs, C. A. (2014). Free-living and particle-associated bacterioplankton in large rivers of the mississippi river basin demonstrate biogeographic patterns. *Appl. Environ. Microbiol.* 80, 7186–7195. doi: 10.1128/aem.01844-14

Kellogg, C. T. E., and Deming, J. W. (2009). Comparison of free-living, suspended particle, and aggregate-associated bacterial and archaeal communities in the Laptev Sea. *Aquat. Microb. Ecol.* 57, 1–18. doi: 10.3354/ame01317

Kim, M., Oh, H., Park, S., and Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 64, 345–351.

Koeppel, A. F., and Wu, M. (2012). Lineage-dependent ecological coherence in bacteria. *FEMS Microbiol. Ecol.* 81, 574–582. doi: 10.1111/j.1574-6941.2012.01387.x

Koeppel, A. F., and Wu, M. (2013). Surprisingly extensive mixed phylogenetic and ecological signals among bacterial operational taxonomic units. *Nucleic Acids Res.* 41, 5175–5188. doi: 10.1093/nar/gkt241

Koeppel, A., Perry, E. B., Sikorski, J., Krizanc, D., Warner, A., Ward, D. M., et al. (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2504–2509. doi: 10.1073/pnas.0712205105

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79, 5112–5120. doi: 10.1128/aem.01043-13

Larkin, A. A., and Martiny, A. C. (2017). Microdiversity shapes the traits, niche space, and biogeography of microbial taxa. *Environ. Microbiol.* 9, 55–70. doi: 10.1111/1758-2229.12523

Lengyel, E., Szabó, B., and Stenger-Kovács, C. (2020). Realized ecological niche-based occupancy-abundance patterns of benthic diatom tratis. *Hydrobiologia* 847, 3115–3127. doi: 10.1007/s10750-020-04324-9

Lindström, E. S., and Langenheder, S. (2012). Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Rep.* 4, 1–9. doi: 10.1111/j.1758-2229.2011.00257.x

Liu, L., Yang, J., Yu, Z., and Wilkinson, D. M. (2015). The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. *ISME J.* 9, 2068–2077. doi: 10.1038/ismej.2015.29

Logares, R., Lindström, E. S., Langenheder, S., Logue, J. B., Paterson, H., Laybourn-Parry, J., et al. (2013). Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME J.* 7, 937–948. doi: 10.1038/ismej.2012.168

Lu, H., Yeh, Y., Sastri, A. R., Shiah, F., Gong, G., and Hsieh, C. (2016). Evaluating community-environment relationships along fine to broad taxonomic resolutions reveals evolutionary forces underlying community assembly. *ISME J.* 10, 2867–2878. doi: 10.1038/ismej.2016.78

Lu, H., Yeh, Y., Shiah, F., Gong, G., and Hsieh, C. (2019). Evolutionary constraints on species diversity in marine bacterioplankton communities. *ISME J.* 13, 1032–1041. doi: 10.1038/s41396-018-0336-1

Mestre, M., Borrull, E., Sala, M. M., and Gasol, J. M. (2017). Patterns of bacterial diversity in the marine planktonic particulate matter continuum. *ISME J.* 11, 999–1010. doi: 10.1038/ismej.2016.166

Milici, M., Vital, M., Tomasch, J., Badewien, T. H., Giebel, H. A., and Plumeier, I. (2017). Diversity and community composition of particle-associated and free-living bacteria in mesopelagic and bathypelagic Southern Ocean water masses: evidence of dispersal limitation in the Bransfield Strait. *Limnol. Oceanogr.* 62, 1080–1095. doi: 10.1002/lno.10487

Mohit, V., Archambault, P., Toupoint, N., and Lovejoy, C. (2014). Phylogenetic differences in attached and free-living bacterial communities in a temperate coastal lagoon during summer, revealed via high-througput 16S rRNA gene sequencing. *Appl. Environ. Microbiol.* 80, 2071–2083. doi: 10.1128/aem.02916-13

Needham, D. M., Sachdeva, R., and Fuhrman, J. A. (2017). Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. *ISME J.* 11, 1614–1629. doi: 10.1038/ismej.2017.29

Nemergut, D. R., Costello, E. K., Hamady, M., Lozupone, C., Jiang, L., Schmidt, S. K., et al. (2011). Global patterns in the biogeography of bacterial taxa. *Environ. Microbiol.* 13, 135–144. doi: 10.1111/j.1462-2920.2010.02315.x

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2020). *Vegan: Community Ecology Package. R Package Version 2.5-7.* Available online at: https://CRAN.R-project.org/package=vegan (accessed January 20, 2021).

Pandit, S. N., Kolasa, J., and Cottenie, K. (2009). Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* 90, 2253–2262. doi: 10.1890/08-0851.1

Philippot, L., Andersson, S. G. E., Battin, T. J., Prosser, J. I., Schimel, J. P., Whiteman, W. B., et al. (2010). The ecological coherence of high bacterial taxonomic ranks. *Nat. Rev. Microbiol.* 8, 523–529. doi: 10.1038/nrmicro2367

Pomeroy, L. R., and Wiebe, W. L. (2001). Temperature and substrates as interactive limiting factors for marine heterotrophic bacteria. *Aquat. Microb. Ecol.* 23, 187–204. doi: 10.3354/ame023187

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). Fasttree: computing large minimum evolution tress with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077

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

Ramachandran, A., and Walsh, D. A. (2015). Investigation of XoxF methanol dehydrogenasesreveals new methylotrophic bacteria in pelagic marineand freshwater ecosystems. *FEMS Microbiol. Ecol.* 91:fiv105. doi: 10.1093/femsec/fiv105

Rieck, A., Herlemann, D. P. R., Jürgens, K., and Grossart, H. (2015). Particle-associated differ from free-living bacteria in surface waters of the Baltic Sea. *Front. Microbiol.* 6:1297.

Ruiz-González, C., Salazar, G., Logares, R., Proia, L., Gasol, J. M., Sabater, S., et al. (2015). Weak coherence in abundance patterns between bacterial classes and their constituent OTUs along a regulated river. *Front. Microbiol.* 6:1293. doi: 10.3389/fmicb.2015.01297

Salazar, G., Cornejo-Castillo, F., Borrull, E., Díez-Vives, C., Lara, E., Vaqué, D., et al. (2015). Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. *Mol. Ecol.* 24, 5692–5706. doi: 10.1111/mec.13419

Saucier, F. J., Roy, F., Gilbert, D., Pellerin, P., and Ritchie, H. (2003). Modeling the formation and circulation processes of water mases and sea ice in the Gulf of St. Lawrence, Canada. *J. Geophys. Res.* 108:3269.

Shade, A., Dunn, R. R., Blowes, S. A., Keil, P., Bohannan, B. J. M., Herrmann, M., et al. (2018). Macroecology to unite all life, large and samll. *Trends Ecol. Evol.* 33, 731–744. doi: 10.1016/j.tree.2018.08.005

Shen, D., Jürgens, K., and Beier, S. (2018a). Experimental insights into the importance of ecologically dissimilar bacteria to community assembly along a salinity gradient. *Environ. Microbiol.* 20, 1170–1184. doi: 10.1111/1462-2920.14059

Shen, D., Langenheder, S., and Jürgens, K. (2018b). Dispersal modifies the diversity and composition of active bacterial communities in response to a salinity disturbance. *Front. Microbiol.* 9:2188. doi: 10.3389/fmicb.2018.02188

Simon, M., Grossart, H. P., Schweitzer, B., and Ploug, H. (2002). Microbial ecology of organic aggregates in aquatic ecosystems. *Aquat. Microb. Ecol.* 28, 175–211. doi: 10.3354/ame028175

Smith, G. C., Saucier, F. J., and Straub, D. (2006a). Formation and circulation of the cold intermediate layer in the Gulf of Saint Lawrence. *J. Geophys. Res.* 111:C06011.

Smith, G. C., Saucier, F. J., and Straub, D. (2006b). Response of the lower St. Lawrence Estuary to external forcing in winter. *J. Phys. Oceanogr.* 36, 1485–1501. doi: 10.1175/jpo2927.1

Smith, M. W., Allen, L. Z., Allen, A. E., Herfort, L., and Simon, H. M. (2013). Contrasting genomic properties of free-living and particle-attached microbial assemblages within a coastal ecosystem. *Front. Microbiol.* 4:120. doi: 10.3389/fmicb.2013.00120

Soininen, J., and Heino, J. (2005). Relationships between local population of species: distribution patterns of diatoms in boreal streams. *J. Biogeogr.* 32, 1971–1978. doi: 10.1111/j.1365-2699.2005.01342.x

Therriault, J. C., and Levasseur, M. (1985). Coltrol of phytoplankton production in the lower St. Lawrence Estuary: light and freshwater runoff. *Naturaliste Can.* 112, 77–96.

Thibodeau, B., de Vernal, A., Hillaire-Marcel, C., and Mucci, A. (2010). Twentieth century warming in deep waters of the Gulf of St. Lawrence: a unique feature of the last millennium. *Geophys. Res. Lett.* 37:L17604.

Van der Gast, C. J., Walker, A. W., Stressmann, F. A., Rogers, G. B., Scott, P., Daniels, T. W., et al. (2011). Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J.* 5, 780–791. doi: 10.1038/ismej.2010.175

Vincent, W. F., and Dodson, J. J. (1999). The St. Lawrence river, Canada-USA: the need for an ecosystem-level understanding of large rivers. *Jpn. J. Limnol.* 60, 29–50. doi: 10.3739/rikusui.60.29

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/aem.00062-07

Wiegand, S., Jogler, M., Boedeker, C., Pinto, D., Vollmers, J., Rivas-Marín, E., et al. (2020). Cultivation and functional characterization of 79 planctomycetes uncovers their unique biology. *Nat. Microbiol.* 5, 126–140. doi: 10.1038/s41564-019-0588-1

Winkler, G., Dodson, J. J., Bertrand, N., Thivierge, D., and Vincent, W. F. (2003). Trophic coupling across the St. Lawrence River estuarine transition zone. *Mar. Ecol. Prog. Ser.* 251, 59–73. doi: 10.3354/meps251059

Yang, Q., Dong, J., Ahmad, M., Ling, J., Zhou, W., Tan, Y., et al. (2019). Analysis of nifH DNA and RNA reveals a disproportionate contribution to nitrogenase activities by rare plankton associated diazotrophs. *BMC Microbiol.* 19:188. doi: 10.1186/s12866-019-1565-9

Zakardjian, B. A., Gratton, Y., and Vézina, A. F. (2000). Late spring phytoplankton bloom in the Lower St. Lawrence Estuary: the flushing hypothesis revisited. *Mar. Ecol. Prog. Ser.* 192, 31–48. doi: 10.3354/meps192031

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., et al. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* 28, 1857–1862. doi: 10.1111/mec.15060