frontiers

Co-association of Bacterial Phylogenetic Diversity with Transparent Exopolymer Particle Distributions in the Surface Microlayer of an Estuary

Nikita Poplavski [1,*]

[1] *School of Natural and Environmental Sciences, Newcastle University, Newcastle-upon-Tyne, UK*

Correspondence:
n.poplavski1@ncl.ac.uk

ABSTRACT

There is growing recognition that the sea-surface microlayer (SML) is subject to significant enrichment with transparent exopolymer particles (TEP), an abundant class of high molecular weight polysaccharides – rendering it to be a highly heterotrophic system. Progressive eutrophication of coastal and estuarine systems may further enhance heterotrophic processes occurring in the SML, reversing it to become a net source of climate-active gases such as CO_2. Yet there is very limited understanding of natural bacterial community dynamics specifically linked to TEP abundances. In this study, coupled SML and cognate underlying water samples from River Tyne (UK) lower estuary and North Sea coastal sites were subjected to high-throughput sequencing of 16S rRNA gene transcripts. Consistent with previous research, Gammaproteobacteria and Bacteroidetes dominated both bacterioplankton and SML communities, although no distinctive niche specialisation even at lower taxonomic resolutions was identified. Parallel analytic TEP assays and flow cytometric cell counts revealed weak responses in community composition, suggesting TEP assimilation may not be unique ecophysiological trait for SML-associated species. The strongest underlying structuring effect was attributed to changes in salinity.

Keywords: sea-surface microlayer, SML, estuary, TEP, metagenomics, bacterioneuston

1 INTRODUCTION

As an interface between gaseous and liquid phases of the two major biotopes – ocean and atmosphere – the (sea-)surface microlayer (SML) is increasingly appreciated as an ecosystem in its own right, possessing profoundly different physiochemical and/or biological properties (Wurl et al., 2017). Due to its unique position at the mass boundary layer, the SML has a great potential to modify air-sea momentum with direct implications on biogeochemical cycles (Wotton and Preston, 2005) and global climate patterns (Liss et al., 2005), for example, via suppression of air-sea gas exchange (Jenkinson et al., 2018). The SML remains intact beyond average open ocean meteorological conditions, above the mean global wind speed of 6.6 ms^{-1} up to at least 13 ms^{-1}, and thus, at any point in time, potentially covers up to 70% of the Earth's surface (Wurl et al., 2011; Sabbaghzadeh et al., 2017). Since in wider literature SML may refer to both the visible monomolecular "slick" (3-4 nm thick) which dampens short gravity-capillary waves, and also its invisible counterpart with a thicker layer of underlying organic matter (OM), within the scope of this work no explicit distinction between the two is made. The SML's depth is loosely considered to lie between 1 and 1000 μm, but more recently complex physical stratification has also been defined: diffusion-limited- ($\leqslant$60 μm), thermal- ($\leqslant$500 μm), and viscous ($\leqslant$1500 μm) sublayers (Hardy, 1982; Zhang et al., 2003; Soloviev et al., 2014). Therefore, no single unified paradigm with standardised SML dimensions is generally accepted (Engel et al., 2017); the rates of geochemical transformations as well as habitat limits for surface-dwelling organisms (collectively called *neuston*), are all to larger extent operationally determined by the ecological feature of interest (Liss et al., 2005).

The SML has been found to selectively accumulate diverse organic and inorganic compounds (van Pinxteren et al., 2012). Lipids, carbohydrates and amino acids/proteins constitute the main dissolved [D]OM pool, whereas coagulation of biogenic polysaccharides forms an abundant class of colloidal microgels and transparent exopolymer particles (TEP) with surface-active (*i.e.* amphiphilic) properties (Verdugo et al., 2004). Owing to the inherent adhesiveness of gel aggregates that enables to become ballasted with other C-rich matter and growth-limiting nutrients, and by virtue of their neutral buoyancy upon formation in the epipelagic, the TEP are widely documented to selectively enrich SML (Wurl and Holmes, 2008; Cunliffe and Murrell, 2009). This has lead to the recognition of SML as a highly hydrated heterogeneous gel polymer matrix (Sieburth, 1983).

In turn, for providing a biofilm-like organic substrate for attachment, the SML hosts enhanced microbial heterotrophic activity (Obernosterer et al., 2005). Besides storing a rich source of bioavailable carbon and nutrients for growth, marine gels, and especially TEP, give refuge from grazing (Mari and Kiørboe, 1996) and detrimental ultraviolet-B solar irradiation (Santos et al., 2012). The composition, relative abundance, and physiology of bacterioneuston assemblages can therefore be significantly different from

the cognate water column communities (Cunliffe et al., 2011). However, studies directly linking *in situ* TEP concentrations to bacterioneuston community diversity using culture-independent methods have not been reported, albeit analogical studies for planktonic assemblages already exist (*e.g.* Busch et al., 2017; Taylor and Cunliffe, 2017).

While the pivotal role of heterotrophic bacteria in OM remineralisation and sustaining the classical planktonic food-web via the microbial loop has attracted considerable research effort (Azam and Malfatti, 2007), most of the knowledge-base in practice was derived from bacterioplankton population surveys. Previous research on estuarine and marine SML systems nevertheless demonstrated patterns in neuston community structure that correlate with environmental factors like salinity (Azevedo et al., 2012), wind and solar forcing (Stolle et al., 2011), hydrodynamic regime (Liparoto et al., 2017), and organic/inorganic matter enrichment (Santos et al., 2013). Cunliffe et al. (2008) has also specifically targeted the functional diversity of bacterioneuston that utilise methane and carbon monoxide, confirming that the habitat filtering by certain taxa is indeed ecologically regulated. The current phylogenetic overview of bacterioneuston diversity consistently recognises four main phyla: Gammaproteobacteria, Bacteroidetes, Actinobacteria and Cyanobacteria (Cunliffe et al., 2011). But at the same coarse taxonomic resolution such ordering is identical to the most abundant planktonic bacterial groups (Barberán and Casamayor, 2010). Considering that the underlying environmental drivers linked to the observed community composition may be to some extent different between SML and the subsurface waters (SSW), an impetus for integrating metagenomes with raw analytical data for SML remains to be satisfied.

As such, the purpose of this study was to apply recent advances in methodological metagenomic techniques (Tikhonov et al., 2015; Pollock et al., 2018) to explore the spatial variability of paired SML and SSW bacterial 16S rRNA clone libraries in response to natural environmental gradients; the latter were typified by a temperate estuarine ecosystem and nearby North Sea coastal sites. Particular focus was set on the potential co-variability of observed TEP concentrations with associated bacterial population metrics, including both the cell abundance and compositional data. Thus, it was hypothesised that (1) the bacterial diversity or abundance are significantly different between the two sampled habitats, and (2) bacterioneuston/-plankton community-wide changes can be adequately approximated as a function of salinity and/or TEP concentration. Finally, bacterial diversity patterns are discussed in the wider context of environmental and physiological attributes linked to potential ecological dynamics and the choice of taxonomic resolution for inferring real biological interactions within the communities.

2 METHODS AND MATERIALS

2.1 Study site

Seawater samples were collected on 18 December 2017 over the course of a single cruise on R/V *Princess Royal* with 5 estuarine stations 1.6 km apart within tidal limits of the River Tyne. One additional station 1 km off the Blyth in coastal North Sea typified a semi-marine system linked to estuarine dynamics due to northward inshore currents (Otto et al., 1990). River Tyne is a highly industrialised waterway in the North East England, subject to numerous point-source inputs of organic matter – including the municipal sewage outfall opposite sampling station 4 – and displaying contrasting SML enrichment profiles from marine (stations 1-2) into brackish (stations 3-6) water characteristics (**Figure 1**).

2.2 Sample Collection

For each respective SML sampling site, bulk subsurface water (SSW) samples were collected in 1 L-triplicates from approximately 0.4 m depth using a bucket to draw water on deck and hand-dipping procedure for collection into pre-sterilised glass bottles (full details in Cunliffe and Wurl, 2014). In addition, metadata from manually deployed CTD probe (U-52G Multiparameter Water Quailty Checker, Horiba Instruments Ltd., Japan) and vessel-based sensors were retrieved for every station (**Table 1**). Immediately upon collection all samples were stored in dark below 4°C and processed within 2 weeks.

A Garrett screen sampler (75 cm × 75 cm) was utilised for SML collection following the recommended protocol by SCOR SML Working Group (2014). For the present study, operationally-defined SML depth congruent with resolution of the chosen sampling technique was estimated to be $440 \pm 120\ \mu$m (S.E.). Sampling was carried out by preliminarily saturating the mesh screen with site-specific seawater constituents via two consecutive 5-second submersions, and upon third submersion, retrieving the screen sampler horizontally from underneath hitherto undisturbed SML. In total, 7-15 surface contacts were carefully funnelled into pre-sterilised borosilicate glass bottles to pool 1 L of SML from each station. Cross-contamination of SML with underlying seawater was minimised by rejecting the first 20 s of drainage from Garrett screen; equipment that has been in use during sampling was sprayed with 70% ethanol between stations and rinsed with site seawater.

2.3 Sample Processing for Downstream Metagenomic Analysis

Both subsurface and SML water samples were serially filtered under low vacuum (ca. -200 mbar) through polycarbonate filters (PC; Whatman Cyclopore) 47 mm in diameter and pore sizes 1.0 μm and 0.2 μm as upper- and lower-bound, respectively. Only the 0.2 μm filter was retained for further processing steps. The maximum volume of aliquots per filter varied between 100-400 ml depending on turbidity of

the sample seawater, but the total volume filtered for each sample equalled 500 ml. The whole procedure was carried out aseptically using autoclaved instruments for handling filters; having processed any single seawater batch, prepared PC membranes were immediately stored away at -20°C in individual sterile 50 ml–centrifuge tubes.

Total DNA extraction from filters comprised of a modified PowerSoil DNA Isolation Kit (MO-BIO, USA), where the standard cell lysis stages were replaced with an alternative workflow combining mechanical disruption, osmotic shock, and chemical lysis. The latter two were an adaptation of the protocol *sensu* Cunliffe et al. (2009). Each centrifuge tube containing a filter was added approximately one-fifth volume of pre-autoclaved 2 mm-diameter glass beads and 2 ml of sucrose lysis buffer (SLB; 750 mM sucrose, 40 mM NaCl, 40 mM EDTA at pH 8.0, 50 mM Tris-HCl at pH 7.5) with 50 μl of Lysozyme solution (final concentration 0.1 mg ml^{-1}) – followed by vigorous vortexing for 1 minute. The tubes were then incubated for 30 minutes at 37 °C with periodic vortexing. Additional 155 μl of 5X Proteinase K, 385 μl of 20% SDS and 562 μl of fresh SLB/Lysozyme solution were added into each tube, briefly vortexed, and placed in a 30 °C incubator with agitation for further 15 hours. After the overnight incubation period, 1 ml of well-mixed lysate was used as per manufacturer-provided instructions for remainder of the procedure. Isolated DNA pellets were consequently air-dried and resuspended in 10X TE buffer to be stored at -20°C.

The quality of DNA samples was verified against negative control extractions via PCR amplification of the 16S rRNA gene using bacterial domain-specific set of primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-CTACGGCTACCTTGTTACGA-3') (Lane, 1991). The PCR master mix was prepared at 0°C on ice and contained 1 μl of both forward and reverse primers (Eurogentec, Belgium), 1 μl DNA, 25 μl MyTaq polymerase (Bioline, UK) and autoclaved MilliQ water to a final volume of 50 μl. The thermocycler programme was set to initial denaturation at 95°C for 3 minutes, followed by 35 cycles of *(1)* 95°C for 30 seconds, *(2)* 55°C for 30 seconds, *(3)* 72°C 30 min, *(4)* and a final elongation step at 72°C for 5 minutes. The PCR amplification results were visualised by agarose gel electrophoresis (40 min at 100 V), whereby 5 μl of PCR product (expected length $\sim$ 1.1 kb) was loaded onto 8% *w/v* agarose gel submerged in TAE 1X buffer (20 mM acetic acid, 40 mM Tris-HCl, 1 mM EDTA) with 3 μl Midori Green nucleic acid stain (Nippon Genetics, Japan). Concentration and purity of DNA was also confirmed with NanoDrop 1000 spectrophotometer (Thermo Scientific, DE) to finalise the selection of a single representative replicate from each sample that will be subject to 16S library preparation ($n = 15$) externally at NU-OMICS facility in the Northumbrian University. Consecutively, paired-end 250 bp sequencing of V3-4 hypervariable region amplicons was performed on the Illumina MiSeq platform and processed via Genome Analyser Pipeline version 1.8 software.

2.4 Colorimetric TEP Assay

The concentration of TEP for each water sample was quantitatively estimated using an empirical relationship between alcian blue staining capability and gel polymer weight as determined with gum xanthan calibration standard (Passow and Alldredge, 1995; Fatibello et al., 2004). Avoiding clogging of the membranes (as observed by sudden decrease in the filtration rate), volume sizes 100-500 ml were filtered onto 0.4 μm PC filters (Whatman Cyclopore) at low vacuum pressure approximately -200 mbar. For each respective sample 4 replicate filters were prepared and frozen at -20°C until further analysis. Subsequently, the thawed filters were stained for 40 seconds with a pre-filtered (0.2 μm pore size) solution of 0.02% *m/v* Alcian Blue 8GX (Merck KGaA, Germany) and 0.6 mol l^{-1} acetic acid (pH 2.5). Excess staining solution was removed by rinsing the stained filters once with MilliQ water, followed by 18 h soaking period in 4 ml of 80% sulphuric acid to dissolve the stain. The resultant solution 1.5 ml aliquots were then centrifuged for 30 s at 10,000 g and absorbances analysed via Cary 100 Bio UV-VIS spectrophotometer (Varian, USA) at 787 nm wavelength in a 1 cm–path cuvette. Reported TEP concentrations were derived from equations given by Passow and Alldredge (1995) and expressed in terms of gum xanthan equivalents (μg l^{-1}).

2.5 Flow Cytometry

Total number of bacteria were analysed in a BD Acurri C6 flow cytometer (BD Biosciences, USA) equipped with 488nm Argon laser. On the next day after collection, untreated water sample volumes of 10 ml were vortexed for 30 seconds with 1% *m/v* SDS to facilitate resuspension of particle-attached bacteria and filtered through 50 μl pore membranes (Whatman Cyclopore). The filtrate was then stained under dark conditions with 3.34 mM SYTO 9 nucleic acid stain (Invitrogen, UK) and incubated at room temperature whilst protected from light for 15 minutes. Stained bacteria fluorescence was collected in a green channel between 515-545 nm and right angle scatter with logarithmic signal amplification (Marie et al., 1997; Marie et al., 2000).

2.6 Data Analysis

All of the data handling and statistical analysis for this study was performed in *R* (v. 3.4.4; R Core Team, 2018) with the modular functionality of packages cited below. The raw sequencing reads were processed using the 'DADA2' pipeline (v. 1.7.7; Callahan et al., 2016): from quality filtering up to constructing an amplicon sequence variants (ASV) table. The ASVs resolve unique sequence features down to single-nucleotide differences as opposed to closed-reference clustering algorithms that pick operational taxonomic units with an arbitrary dissimilarity threshold (e.g. UCLUST, mothur)(Callahan et al., 2017). Taxonomic classifications were therefore assigned to genus-species by exact matching using the RDP's naive Bayesian classifier trained on Silva taxonomic reference data (release 132). Chimeric sequences, accounting for

$< 2\%$ of total sequence reads, were removed with the package 'DECIPHER' (v. 2.6.0), which was also used to perform multiple sequence alignment for subsequent phylogenetic tree reconstruction(Wright, 2016). The latter was realised within the package 'phangorn' (v. 2.4.0) via pairwise distance-matrix calculations based on maximum likelihood methods and estimating tree structure with neighbour-joining and UPGMA algorithms (Schliep et al., 2017). Final quality screening step consisted of rarefaction analysis on raw ASV abundance counts to assess the sampling effort per sample.

Observed richness, Shannon and Simpson indices were calculated to infer within-sample (α-) diversity across samples. As a between-sample (β-) diversity measure, pairwise differences between pooled SML and SSW indices were compared using Wilcoxon signed-rank test. The bacterial community profiles were examined via hierarchical clustering on Jensen-Shannon Divergence (JSD) distance matrix in package 'pvclust' (v. 2.0), yielding p-values for each of the clusters that were then incorporated into a dendrogram (Suzuki and Shimodaira, 2006). To further explore spatial and environmental forcing effects, a series multivariate ordination analyses with inherently different assumptions (and hence robustness/sensitivities) was conducted (NMDS, DPCoA, CCA); packages 'phyloseq' (v. 1.22.3; McMurdie and Holmes, 2013) and 'vegan' (v. 2.4.6; Dixon, 2003) were utilised for these and the following statistical methods. Overall differences in SML and SSW taxonomic assemblages were assessed in the analysis of similarities (ANOSIM) based on Bray-Curtis and abundance-weighted UniFrac distance matrices with 999 permutations. Prior to constructing a Bray-Curtis similarity matrix, the ASV abundance data was square-root transformed to weaken heavy tails of rare and very abundant taxa. To verify the extent of any ANOSIM significance as a product of dispersion within samples, a non-parametric permutational analysis of variance (PERMANOVA) was performed in parallel with same parameters as for ANOSIM.

3 RESULTS

3.1 Sequencing Overview

In total 1,732,795 short-read sequences were compiled into a 16S rRNA clone library for this study (**Table 2**). Quality-filtering reduced the number of raw sequences down to 1,620,907, on average $-8.14\pm4.71\%$ (S.D.) per sample, which in turn revealed 3,443 ASVs across 15 environmental samples. Adequate sequencing depth was evidenced by all of the rarefaction curves plateauing out beyond exponential phase, suggesting that the recovered metagenomes constrained most of the environmental phylotypes and could be thus considered representative (**Figure 2**).

3.2 Bacterial Diversity

Among 6 sampling sites, α-diversity indices were found to be the highest for SML upper-estuarine sites 5 and 6 (salinities 19.5 and 9.5, respectively)(with exception of SSW-6), whereas the lowest diversity was affiliated with another SML sample at the opposite survey range, marine site 6 ($S = 32.5$)(**Figure 3**). An inverse β-diversity pattern was generally observed in subsurface-derived samples: SSW index maxima at site 1/2 transitioned to a uniform minima in both Shannon and Simpson diversity interpretations at site 5. Despite samples SML-3 and SSW-6 contained two replicates each (unlike the rest with single amplicon libraries), the latter had abnormally high α-diversity above or comparable to SML-6. Considering that the same peak was consistently produced with 2 independent observations, SSW-6 data were retained for further analysis instead of discarded as outliers. Following the Wilcoxon signed-rank test on pooled α-diversity scores for SML and SSW, no significant differences ($p > 0.05$) were found for raw species richness distributions. However, if taking into account species evenness, the opposite was true for both used metrics ($p < 0.05$)(**Figure 4**). In comparison with Shannon index, Simpson scores are more weighted towards species evenness, so higher inferred significance for the latter ($p > 0.01$) suggested species evenness as the principal causal component explaining differences between communities.

According to Silva 16S reference database, overall 51 phyla were identified, of which 5 phyla across 20 genera covered $> 92\%$ of the sequences (**Figure 5**). Community composition was heavily dominated by class Gammaproteobacteria, contributing on average more than half of the total taxonomic diversity. The second- and third-ranked omnipresent taxa belonged to phylum Bacteroidetes and Epsilonbacteraeota (in some other reference databases usually assigned to Alphaproteobacteria), respectively, with relative abundances generally below 10% but sporadically peaking above 30%, e.g. SML-4 and SSW-5 (**Figure 6**). Members of the genus *Glutamibacter* (Actinobacteria), among several other singleton clades which entailed only 1 lowest taxonomic-rank member, were most likely a result of contamination during sample handling. On the contrary, a sharp peak of unidentified Cyanobacteria (order Enterobacteriales) group was seen in SSW-3 sample, but no significant signal was present in control filter library. Because the potential taxonomic diversity of control sequences as well as total cell abundance (**Table 1**) were comparatively low, the potential bias introduced from contamination was assumed to be insignificant relative to real fine-scale biological variation. However, apparent systematic differences in taxonomic composition between individual sites nor bacterial assemblages from paired samples were not identified.

A similar conclusion was drawn from hierarchical clustering analysis, where clusters tended to occur on identical sites and branching was to larger extent an artefact of intra-sample dispersion rather than a result of habitat structuring (**Figure 7**). Nonetheless, with caution, it may be stated that a distinct marine cluster at 66% identity was formed (but see SSW-5).

3.3 Environmental Forcing of Community Structure

The variation in phylogenetic β-diveristy using unconstrained ordination methods revealed a consistent ecological gradient of change in taxonomic composition with salinity (**Figure 8**). A weaker community-level response to TEP concentration was also evident from sorting of the individual site data points along the second axis, but absence of expected linearity in taxonomic clustering (i.e. observed formation of four non-monotonic taxonomic lineages) possibly indicated on strong influence of abundance outliers to which DPCoA tends to be more sensitive than other approaches (**Figure 9**). At the collective resolution, emphasizing total variation among individual samples, general tendency towards the formation of two disaligned clusters was observed for SML- and SSW-derived assemblages (**Figure 10**). This implied formation of distinct bacterioneuston and bacterioplankton communities under non-parametric assumptions shared between NMDS and DPCoA ordination methods: namely undefined number of populations and monotonic response function. In addition, comparing the two aforementioned cluster centroid positions using ANOSIM provided some statistical validity using both Bray-Curtis ($p = 0.032$, $R = 0.222$) and unweighted UniFrac ($p = 0.082$, $R = 0.1844$) distance matrices under relaxed 0.1-significance level to account for inconsistent α-diversity patterns. However, between-treatment statistical inference detected significant variability between paired sites ([w-UniFrac] $p = 0.056$, $F = 5.99$; [Bray] $p = 0.068$, $F = 5.99$, PERMANOVA), leading to accept the null hypothesis that the SML and SSW bacterial community structure, regardless of the site, was not significantly different, where positive ANOSIM outcome was mainly impacted by multivariate dispersion. This is further supported by studying the ANOSIM R statistics that represent the difference of mean ranks between groups and within groups; as the reported values tended to approach 0 – indicating no separation on a scale up to 1 (complete separation) – these results are consistent with PERMANOVA and visualised NMDS cluster variances (Figure 8).

4 DISCUSSION

Transition zones in estuarine and coastal areas are characterised by marked shifts in salinity, nutrients, temperature, and pH, among other possible abiotic factors which are central to the spatial variability of planktonic microbial communities (Herlemann et al., 2011). Thus, at the seawater–freshwater interface community-wide responses to environmental fluctuations are likely to be the most exaggerated (Fortunato et al., 2012). In this study, both neustonic and planktonic bacterial community diversities were assessed along the lower-reach River Tyne transect that would span the steepest gradient in OM and TEP enrichment. The observed TEP enrichment factors (EF), calculated as the ratio of analyte concentration in SML with respect to SSW, supported previously reported patterns as the EF for all 3 freshwater/brackish sites was >1 ($\bar{X} = 1.52$) and <1 ($\bar{X} = 0.75$) for sites with larger marine component; what relates to typical

particulate polysaccharide EFs ranging between 1.7 and 7.0 in estuarine SML environments worldwide (Gao et al., 2012; Mari et al., 2012). Interestingly, however, cell abundance data did not correlate with TEP concentrations ($r = 0.39$, n = 14, $p > 0.05$) contrary to general assumption that the availability of carbon-rich substrate favours bacterial colonisation. Similarly, Santos et al. (2013) found only 1.3-fold statistically insignificant EF of β-glucosidase hydrolytic activity in the SML of a brackish Ria de Aveiro estuary zone – that is, compared to a 7-fold higher carbohydrate-hydrolysis enzyme concentrations at surveyed marine sites. A common feature of temperate Northern Hemisphere estuarine systems is the predominantly allochthonous source of OM as opposed to autochthonous processes that will increasingly dominate the total brackish-to-marine OM pool and, importantly, the chemical composition of TEP (Almeida et al., 2002). With riverine DOM being more heterogeneous, it is therefore possible that the resultant TEP may not be as labile for a significant proportion of specialist bacteria.

In a seminal study by Taylor and Cunliffe (2017) to gain a first insight into natural bacterioplankton assemblages utilising diatom-derived TEP, the orders Flavobacteriales and Alteromondales were experimentally demonstrated to directly assimilate TEP, whereas the order Rhodobacteriales hinted at a possible syntrophic link with the former taxa via consumption of secondary monomeric TEP degradation products (Buchan et al., 2014). Based on the high-throughput sequencing results in this work, Alteromondales (Gammaproteobacteria) were found to constitute 63% of the total sequences. It is notable that extremely high abundances of alteromonads in SML-1 and SML-3 samples coincided with nearby urban effluent discharge sources (Figure 5). Alteromondales are known to elicit rapid responses to substrate availability under several DOC-addition mesocosms, displaying them as highly efficient *r*-selected species (Sarmento and Gasol, 2012; Pedler et al., 2014). This may explain observed taxonomic abundance peaks exclusively at the two aforementioned SML samples whereby elevated TEP coagulation rates potentially induced increased upward transport flux .

A generally accepted view that the bacterioneuston assemblages form by recruitment from underlying water column is consistent with the apparent similarity in both community composition profiles (Lindroos et al., 2011). The present work showed that there are no significant β-diversity differences between bacterioneuston and bacterioplankton assemblages if to account for underlying phylogenetic relationships implemented in multivariate ordination methods. Also the statistically significant differences obtained with α-diversity indices were in fact found to be compromised by intra-sample variances as they only consider raw taxonomic counts (Figure 4). From the ecological perspective, bacterial biomass production rates in the SML have been previously shown to be approximately 10 times lower than in the SSW (Santos et al., 2011); short residence times of most estuarine systems due to turbulent hydrodynamic regimes (*e.g.* tidal mixing coupled with shallow bathymetry) prevent establishment of a successional bacterial community within the SML.

REFERENCES

Almeida, M., Cunha, M., and Alcântara, F. (2002). Seasonal change in the proportion of bacterial and phytoplankton production along a salinity gradient in a shallow estuary. *Hydrobiologia* 475, 251–262

Azam, F. and Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology* 5, 782

Azevedo, J. S., Ramos, I., Araújo, S., Oliveira, C. S., Correia, A., and Henriques, I. S. (2012). Spatial and temporal analysis of estuarine bacterioneuston and bacterioplankton using culture-dependent and culture-independent methodologies. *Antonie van Leeuwenhoek* 101, 819–835

Barberán, A. and Casamayor, E. O. (2010). Global phylogenetic community structure and β-diversity patterns in surface bacterioplankton metacommunities. *Aquatic Microbial Ecology* 59, 1–10

Buchan, A., LeCleir, G. R., Gulvik, C. A., and González, J. M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature Reviews Microbiology* 12, 686

Busch, K., Endres, S., Iversen, M. H., Michels, J., Nöthig, E.-M., and Engel, A. (2017). Bacterial colonization and vertical distribution of marine gel particles (TEP and CSP) in the arctic Fram Strait. *Frontiers in Marine Science* 4, 166

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal* 11, 2639

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* 13, 581–583

Cunliffe, M. and Murrell, J. C. (2009). The sea-surface microlayer is a gelatinous biofilm. *The ISME Journal* 3, 1001

Cunliffe, M., Salter, M., Mann, P. J., Whiteley, A. S., Upstill-Goddard, R. C., and Murrell, J. C. (2009). Dissolved organic carbon and bacterial populations in the gelatinous surface microlayer of a Norwegian fjord mesocosm. *FEMS microbiology letters* 299, 248–254

Cunliffe, M., Schäfer, H., Harrison, E., Cleave, S., Upstill-Goddard, R., and Murrell, J. C. (2008). Phylogenetic and functional gene analysis of the bacterial and archaeal communities associated with the surface microlayer of an estuary. *The ISME journal* 2, 776

Cunliffe, M., Upstill-Goddard, R. C., and Murrell, J. C. (2011). Microbiology of aquatic surface microlayers. *FEMS microbiology reviews* 35, 233–246

Cunliffe, M. and Wurl, O. (2014). *Guide to best practices to study the ocean's surface.* SCOR

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* 14, 927–930

Engel, A., Bange, H. W., Cunliffe, M., Burrows, S. M., Friedrichs, G., Galgani, L., et al. (2017). The ocean's vital skin: Toward an integrated understanding of the sea surface microlayer. *Frontiers in Marine Science* 4, 165

Fatibello, S. H. A., Vieira, A. A. H., and Fatibello-Filho, O. (2004). A rapid spectrophotometric method for the determination of transparent exopolymer particles (TEP) in freshwater. *Talanta* 62, 81–85

Fortunato, C. S., Herfort, L., Zuber, P., Baptista, A. M., and Crump, B. C. (2012). Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. *The ISME Journal* 6, 554

Gao, Q., Leck, C., Rauschenberg, C., and Matrai, P. A. (2012). On the chemical dynamics of extracellular polysaccharides in the high Arctic surface microlayer. *Ocean Science* 8, 401–418

Hardy, J. T. (1982). The sea surface microlayer: biology, chemistry and anthropogenic enrichment. *Progress in Oceanography* 11, 307–328

Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., and Andersson, A. F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME journal* 5, 1571

Jenkinson, I. R., Laurent, S., Ding, H., and Elias, F. (2018). Biological modification of mechanical properties of the sea surface microlayer, influencing waves, ripples, foam and air-sea fluxes. *Elem Sci Anth* 6

Lane, D. J. (1991). 16S/23S rRNA Sequencing. In *Nucleic Acid Techniques in Bacterial Systematic*, eds. E. Stackebrandt and M. Goodfellow (New York: John Wiley and Sons). 115–175

Lindroos, A., Szabo, H. M., Nikinmaa, M., and Leskinen, P. (2011). Comparison of sea surface microlayer and subsurface water bacterial communities in the Baltic Sea. *Aquatic Microbial Ecology* 65, 29–42

Liparoto, A., Mancinelli, G., and Belmonte, G. (2017). Spatial variation in biodiversity patterns of neuston in the Western Mediterranean and Southern Adriatic Seas. *Journal of Sea Research* 129, 12–21

Liss, P. S., Liss, P. S., and Duce, R. A. (2005). *The sea surface and global change* (Cambridge University Press)

Mari, X. and Kiørboe, T. (1996). Abundance, size distribution and bacterial colonization of transparent exopolymeric particles (TEP) during spring in the Kattegat. *Journal of Plankton Research* 18, 969–986

Mari, X., Torréton, J.-P., Trinh, C. B.-T., Bouvier, T., Van Thuoc, C., Lefebvre, J.-P., et al. (2012). Aggregation dynamics along a salinity gradient in the Bach Dang estuary, North Vietnam. *Estuarine, Coastal and Shelf Science* 96, 151–158

Marie, D., Partensky, F., Jacquet, S., and Vaulot, D. (1997). Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. *Applied and environmental microbiology* 63, 186–193

Marie, D., Simon, N., Guillou, L., Partensky, F., and Vaulot, D. (2000). Flow cytometry analysis of marine picoplankton. In *In Living Color* (Springer). 421–454

McMurdie, P. J. and Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8, e61217

Obernosterer, I., Catala, P., Reinthaler, T., Herndl, G. J., and Lebaron, P. (2005). Enhanced heterotrophic activity in the surface microlayer of the Mediterranean Sea. *Aquatic Microbial Ecology* 39, 293–302

Otto, L., Zimmerman, J., Furnes, G., Mork, M., Saetre, R., and Becker, G. (1990). Review of the physical oceanography of the North Sea. *Netherlands journal of sea research* 26, 161–238

Passow, U. and Alldredge, A. (1995). A dye-binding assay for the spectrophotometric measurement of transparent exopolymer particles (TEP). *Limnology and Oceanography* 40, 1326–1335

Pedler, B. E., Aluwihare, L. I., and Azam, F. (2014). Single bacterial strain capable of significant contribution to carbon cycling in the surface ocean. *Proceedings of the National Academy of Sciences* 111, 7202–7207

Pollock, J., Glendinning, L., Wisedchanwet, T., and Watson, M. (2018). The madness of microbiome: Attempting to find consensus "best practice" for 16S microbiome studies. *Applied and environmental microbiology* 84, e02627–17

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria

Sabbaghzadeh, B., Upstill-Goddard, R., Beale, R., Pereira, R., and Nightingale, P. (2017). The Atlantic Ocean surface microlayer from 50° N to 50° S is ubiquitously enriched in surfactants at wind speeds up to 13 m s- 1. *Geophysical Research Letters* 44, 2852–2858

Santos, A. L., Baptista, I., Lopes, S., Henriques, I., Gomes, N. C., Almeida, A., et al. (2012). The UV responses of bacterioneuston and bacterioplankton isolates depend on the physiological condition and involve a metabolic shift. *FEMS microbiology ecology* 80, 646–658

Santos, L., Santos, A. L., Coelho, F. J., Gomes, N. C. M., Dias, J. M., Cunha, Â., et al. (2011). Relation between bacterial activity in the surface microlayer and estuarine hydrodynamics. *FEMS microbiology ecology* 77, 636–646

Santos, L., Santos, A. L., Coelho, F. J., Marcial Gomes, N. C., Dias, J. M., Cunha, Â., et al. (2013). Heterotrophic activities of neustonic and planktonic bacterial communities in an estuarine environment (Ria de Aveiro). *Journal of plankton research* 36, 230–242

Sarmento, H. and Gasol, J. M. (2012). Use of phytoplankton-derived dissolved organic carbon by different types of bacterioplankton. *Environmental microbiology* 14, 2348–2360

Schliep, Klaus, Potts, J., A., Morrison, A., D., et al. (2017). Intertwining phylogenetic trees and networks. *Methods in Ecology and Evolution* 8, 1212–1220

Sieburth, J. M. (1983). Microbiological and organic-chemical processes in the surface and mixed layers. In *Air-sea exchange of gases and particles* (Springer, Dordrecht). 121–172

Soloviev, A. V., Lukas, R., Donelan, M. A., Haus, B. K., and Ginis, I. (2014). The air-sea interface and surface stress under tropical cyclones. *Scientific reports* 4, 5306

Stolle, C., Labrenz, M., Meeske, C., and Jürgens, K. (2011). Bacterioneuston community structure in the southern baltic sea and its dependence on meteorological conditions. *Applied and environmental microbiology* 77, 3726–3733

Suzuki, R. and Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22, 1540–1542

Taylor, J. D. and Cunliffe, M. (2017). Coastal bacterioplankton community response to diatom-derived polysaccharide microgels. *Environmental microbiology reports* 9, 151–157

Tikhonov, M., Leach, R. W., and Wingreen, N. S. (2015). Interpreting 16S metagenomic data without clustering to achieve sub-OTU resolution. *The ISME journal* 9, 68

van Pinxteren, M., Muller, C., Iinuma, Y., Stolle, C., and Herrmann, H. (2012). Chemical characterization of dissolved organic compounds from coastal sea surface microlayers (Baltic Sea, Germany). *Environmental science & technology* 46, 10455–10462

Verdugo, P., Alldredge, A. L., Azam, F., Kirchman, D. L., Passow, U., and Santschi, P. H. (2004). The oceanic gel phase: a bridge in the DOM–POM continuum. *Marine Chemistry* 92, 67–85

Wotton, R. S. and Preston, T. M. (2005). Surface films: Areas of water bodies that are often overlooked. *BioScience* 55, 137–145

Wright, E. S. (2016). Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. *The R Journal* 8, 352–359

Wurl, O., Ekau, W., Landing, W. M., and Zappa, C. J. (2017). Sea surface microlayer in a changing ocean–A perspective. *Elem Sci Anth* 5

Wurl, O. and Holmes, M. (2008). The gelatinous nature of the sea-surface microlayer. *Marine Chemistry* 110, 89–97

Wurl, O., Wurl, E., Miller, L., Johnson, K., and Vagle, S. (2011). Formation and global distribution of sea-surface microlayers. *Biogeosciences* 8, 121–135

Zhang, Z., Cai, W., Liu, L., Liu, C., and Chen, F. (2003). Direct determination of thickness of sea surface microlayer using a pH microelectrode at original location. *Science in China Series B: Chemistry* 46, 339–351

FIGURES AND TABLES

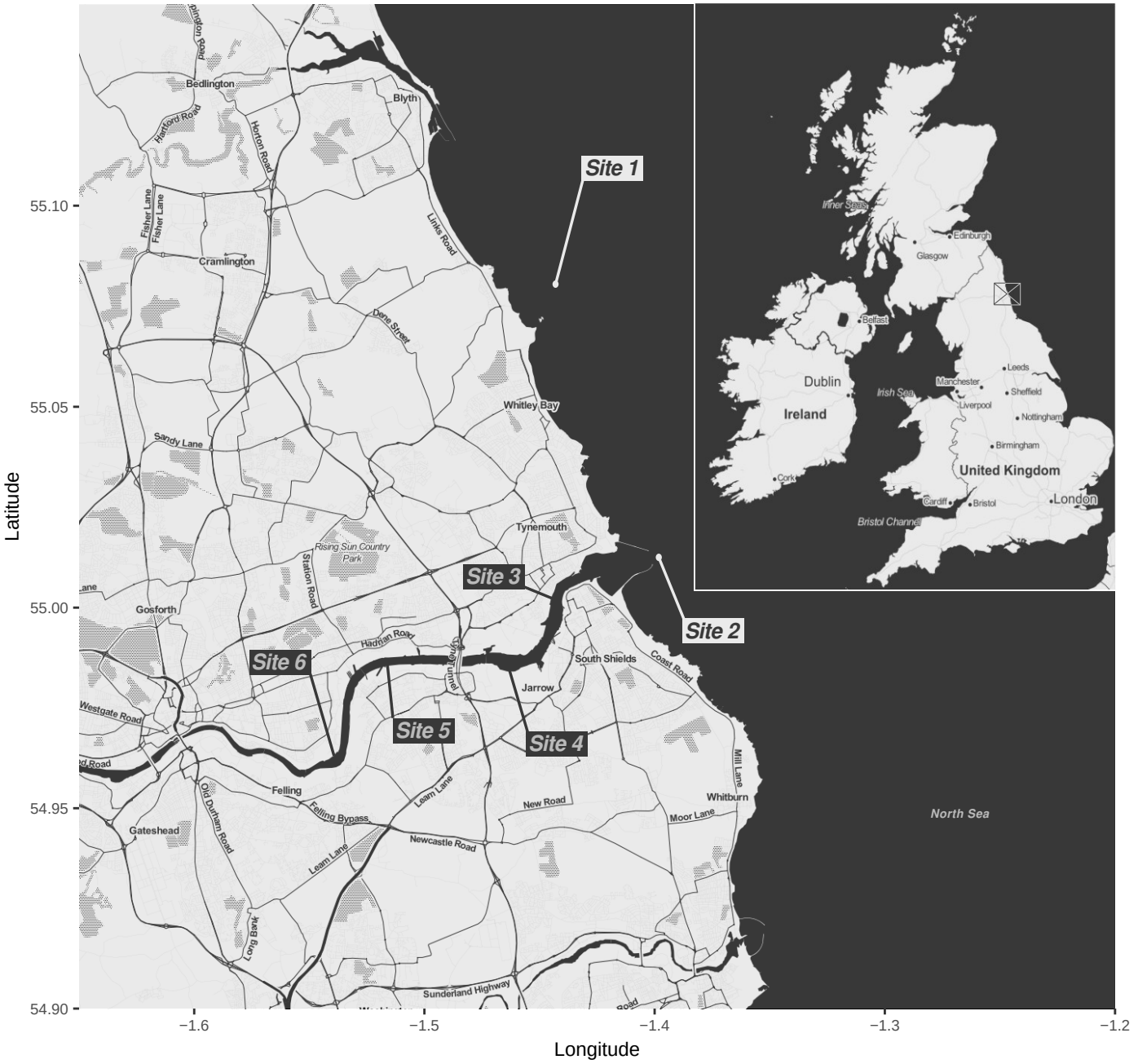

Figure 1. Study location and sampling site mappings.

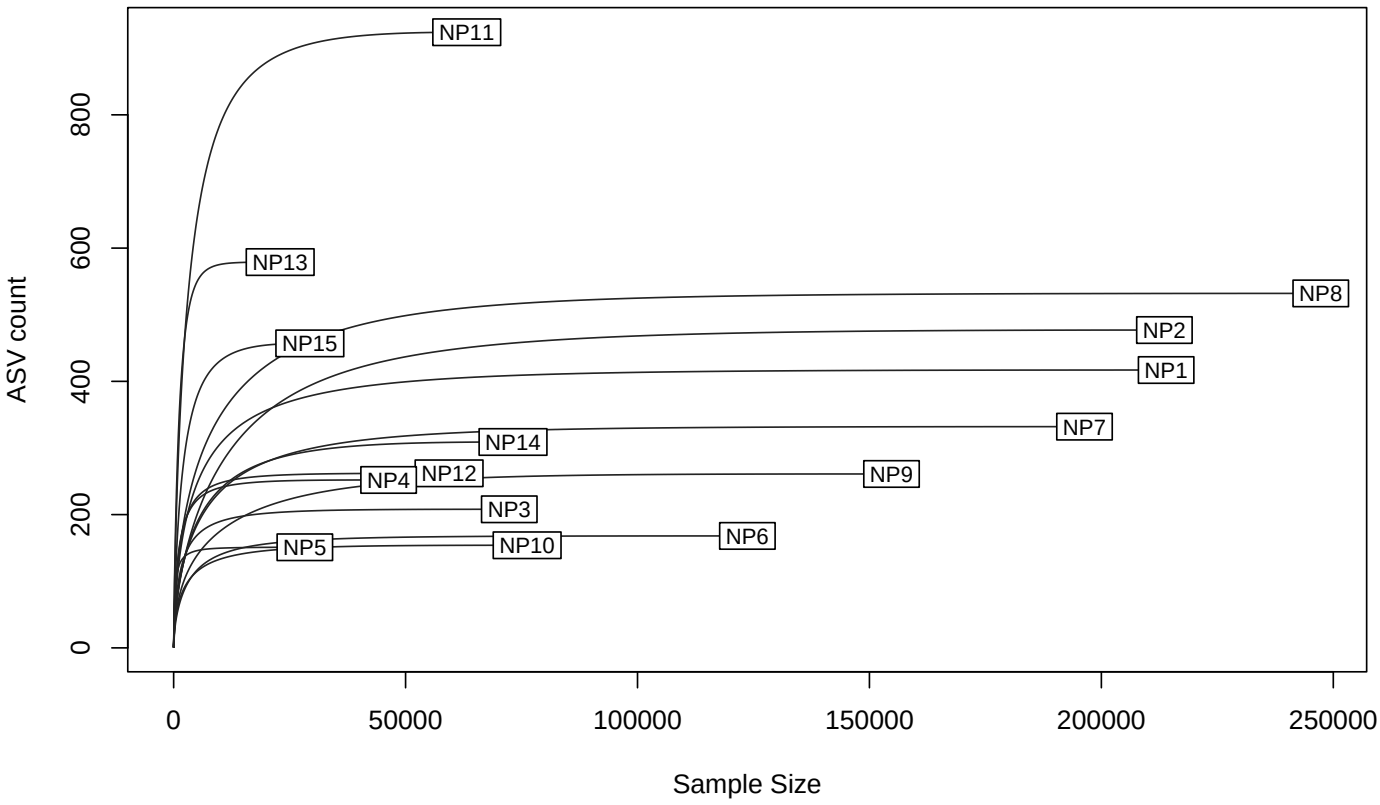

Figure 2. Rarefaction analysis of sampling effort across individual samples. The slope gradient is representative of the probability to discover new taxa when systematically increasing sample size from 0 onwards (step size = 200).

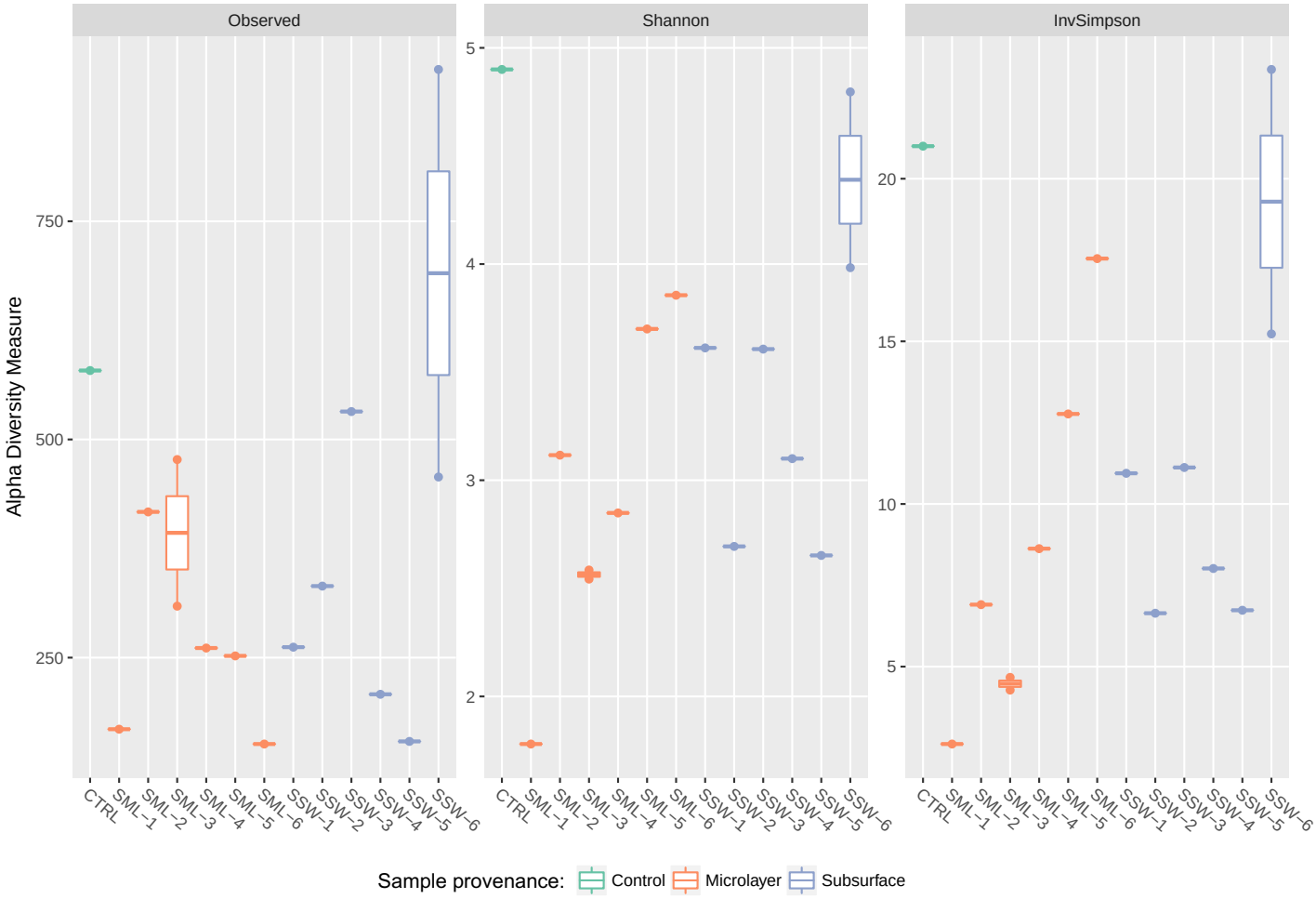

Figure 3. Comparative alpha-diversity measures using raw observations and Shannon-Wiener/Simpson indices.

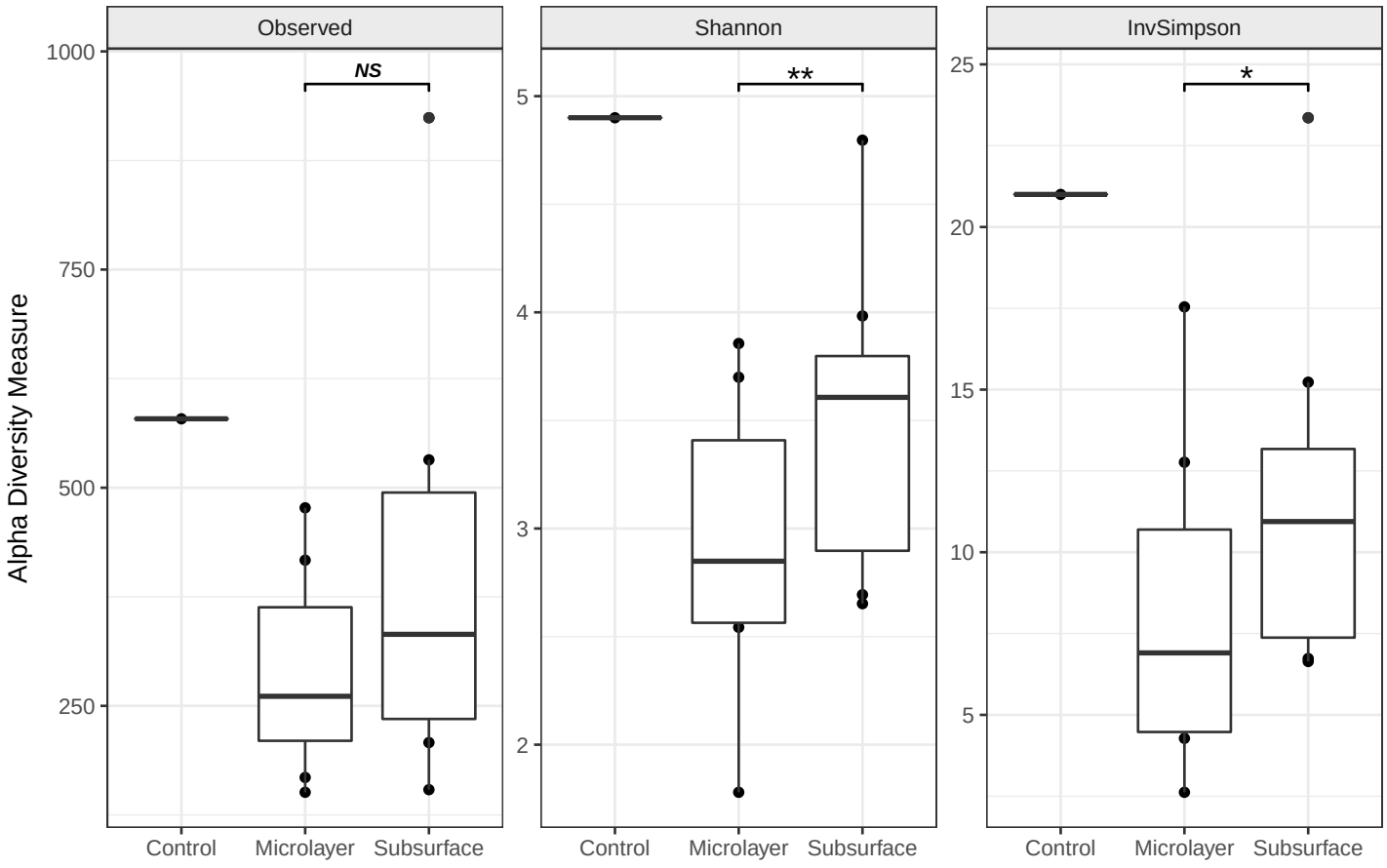

Figure 4. Alpha-diversity measures using raw observations and Shannon-Wiener/Simpson indices. The horizontal bracket above SML and SSW groups displays the results of Wilcoxon signed-rank test, where *NS* stands for 'no significant difference' ($p > 0.05$); * for $p < 0.05$; and ** for $p < 0.01$.

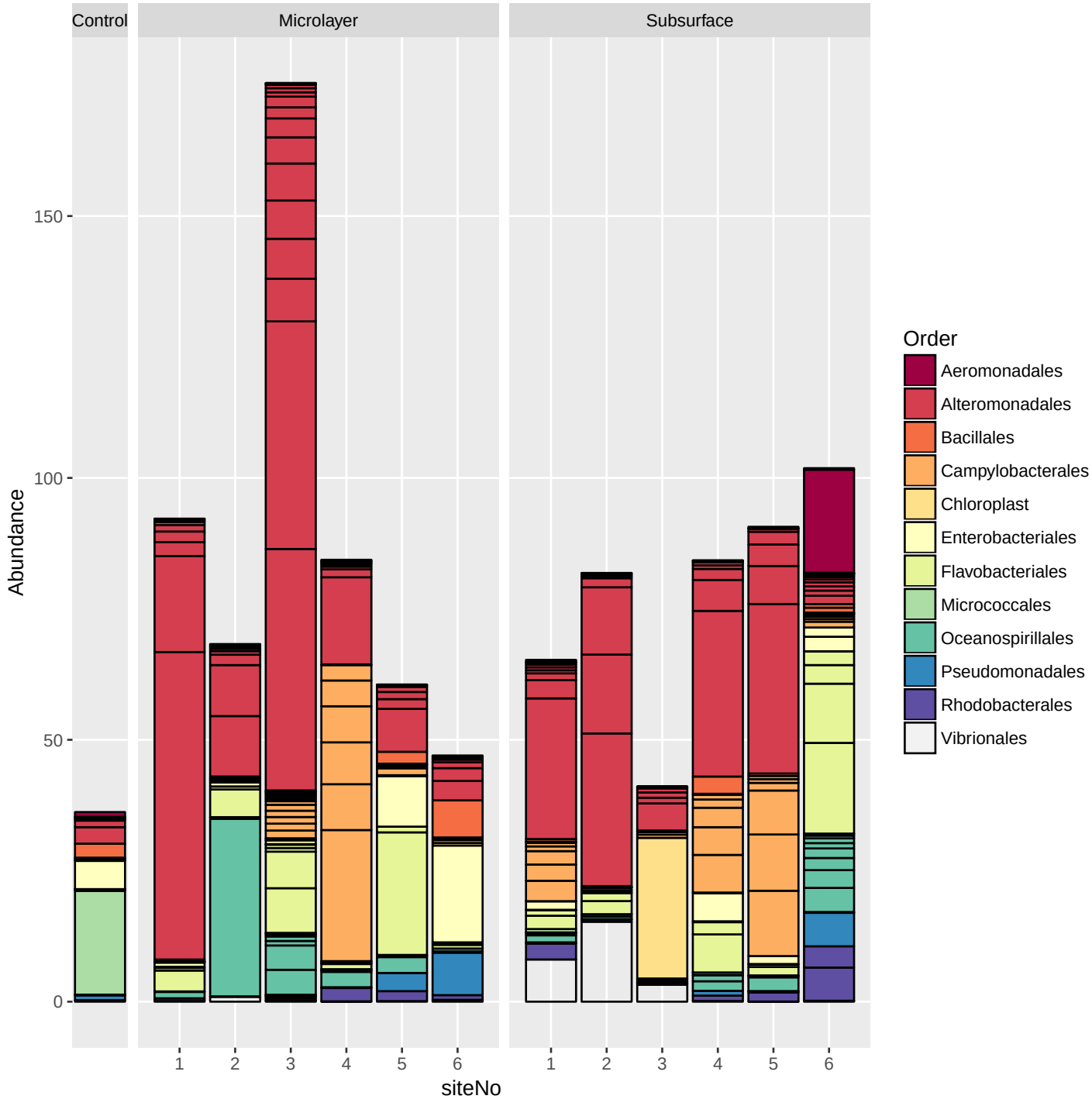

Figure 5. An overview of bacterioneuston (SML) and bacterioplankton (SSW) phylogenetic diversity; stacked segments indicate intra-sample relative taxa abundance grouped at a family level. For the purposes of clarity only the top 20 most common genera are annotated.

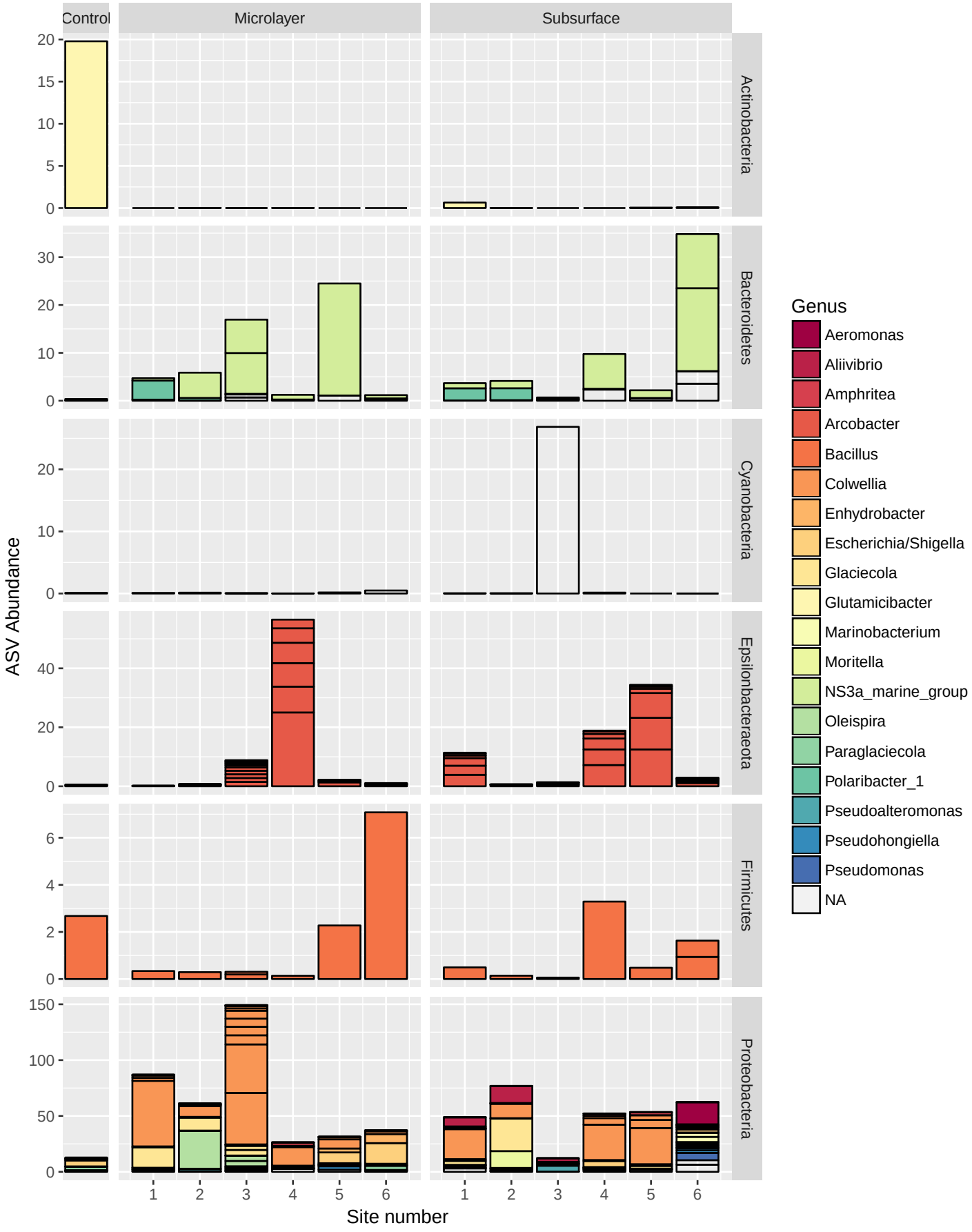

Figure 6. Bacterial community composition in surface microlayer and the cognant subsurface water samples at a genus level with modulation across individual phyla. For the purposes of clarity only the top 20 most common genera are annotated.

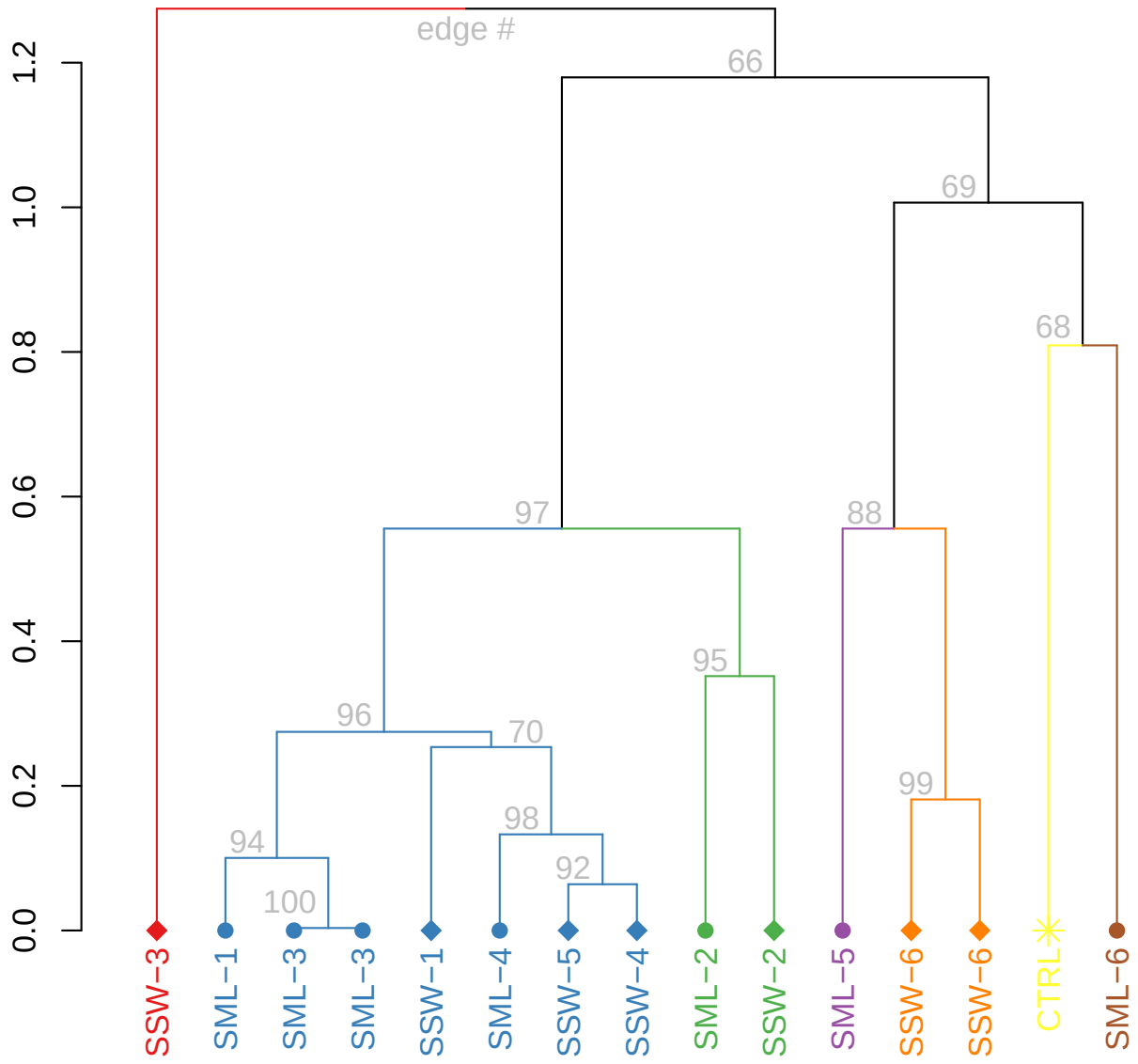

Figure 7. Hierarchical cluster analysis between individual sites and paired sampling environments denoted as SML-/SSW-[siteNo]. Numbers at the branching points give percentage similarity (%) based on Janson-Shannon Divergence distance metric (*y*-axis).

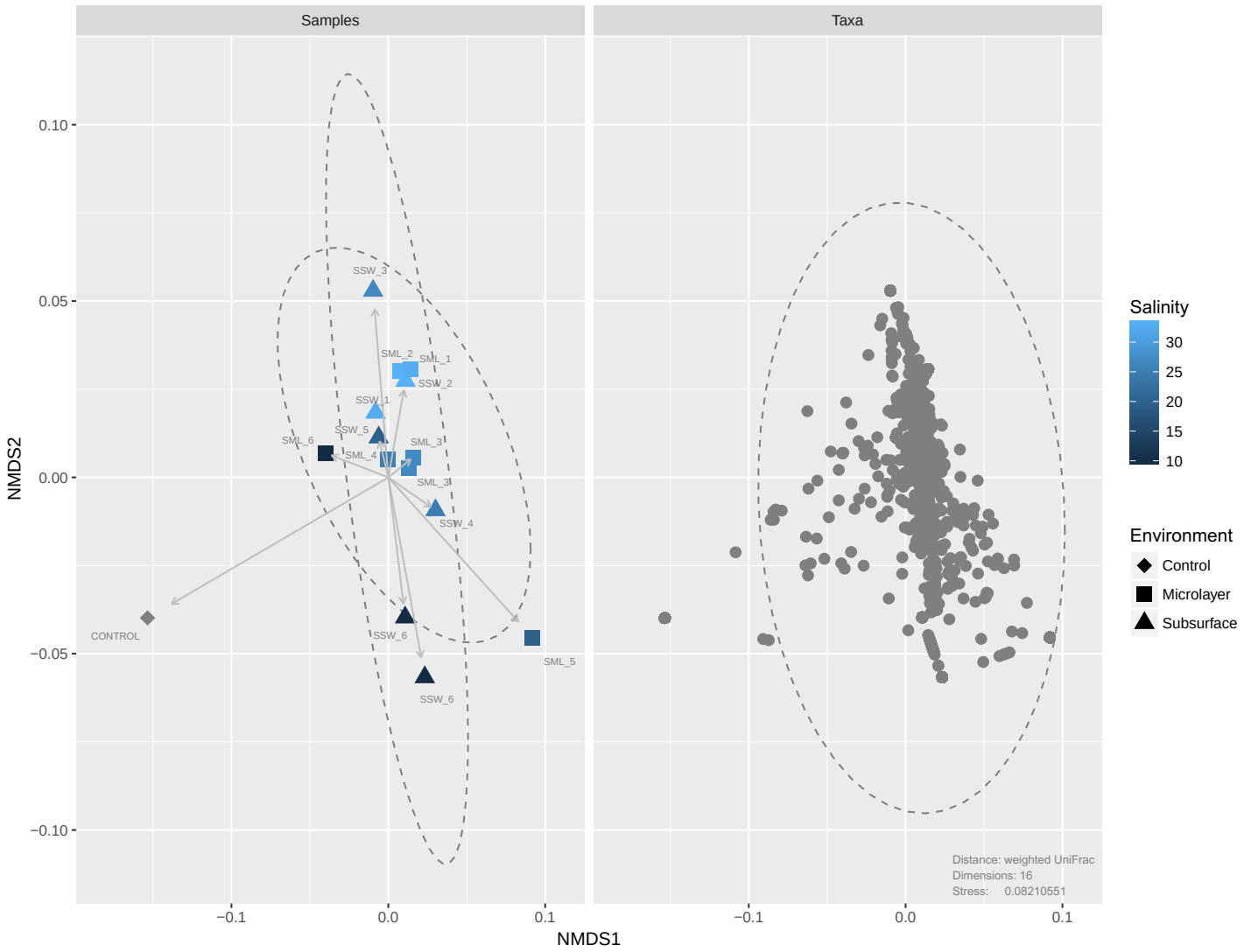

Figure 8. Non-metric multidimensional scaling (NMDS) biplot based on phylogeny-aware weighted UniFrac matrices for individual samples ($n = 15$) in the left panel and pooled taxa to the right. Colour-coded is the effect of water salinity on the multivariate structure of sampling environment. Ellipses indicate approximate clusters for grouping variables and arrow length represent magnitude of the effect on response variables.

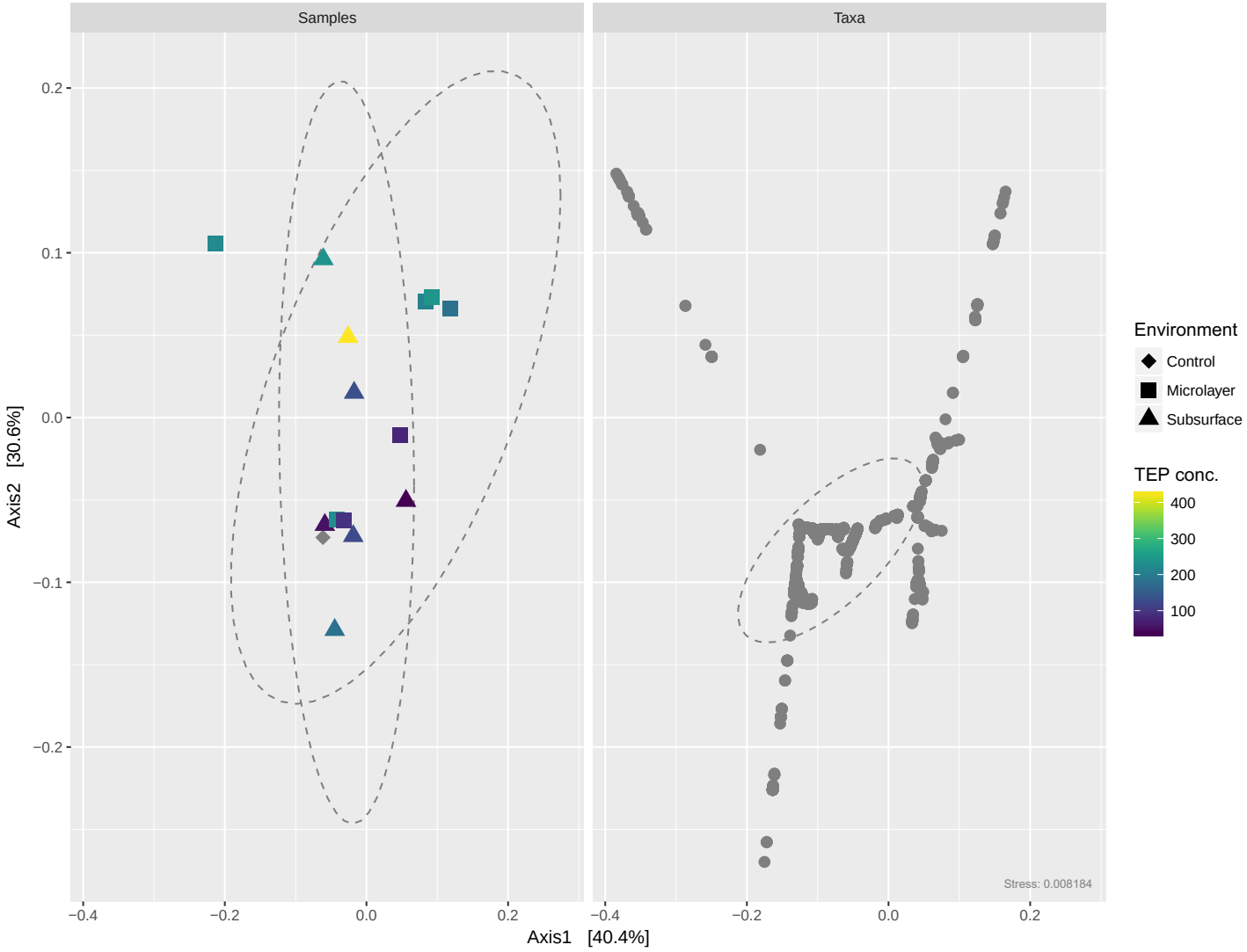

Figure 9. Double principle coordinates analysis (DPCoA) biplot for individual samples ($n = 15$) in the left panel and pooled taxa to the right. Colour-coded is the effect of TEP ($\mu g\, l^{-1}$ gum xanthan equivalent) on the multivariate structure of sampling environment. Ellipses indicate approximate clusters for grouping variables.

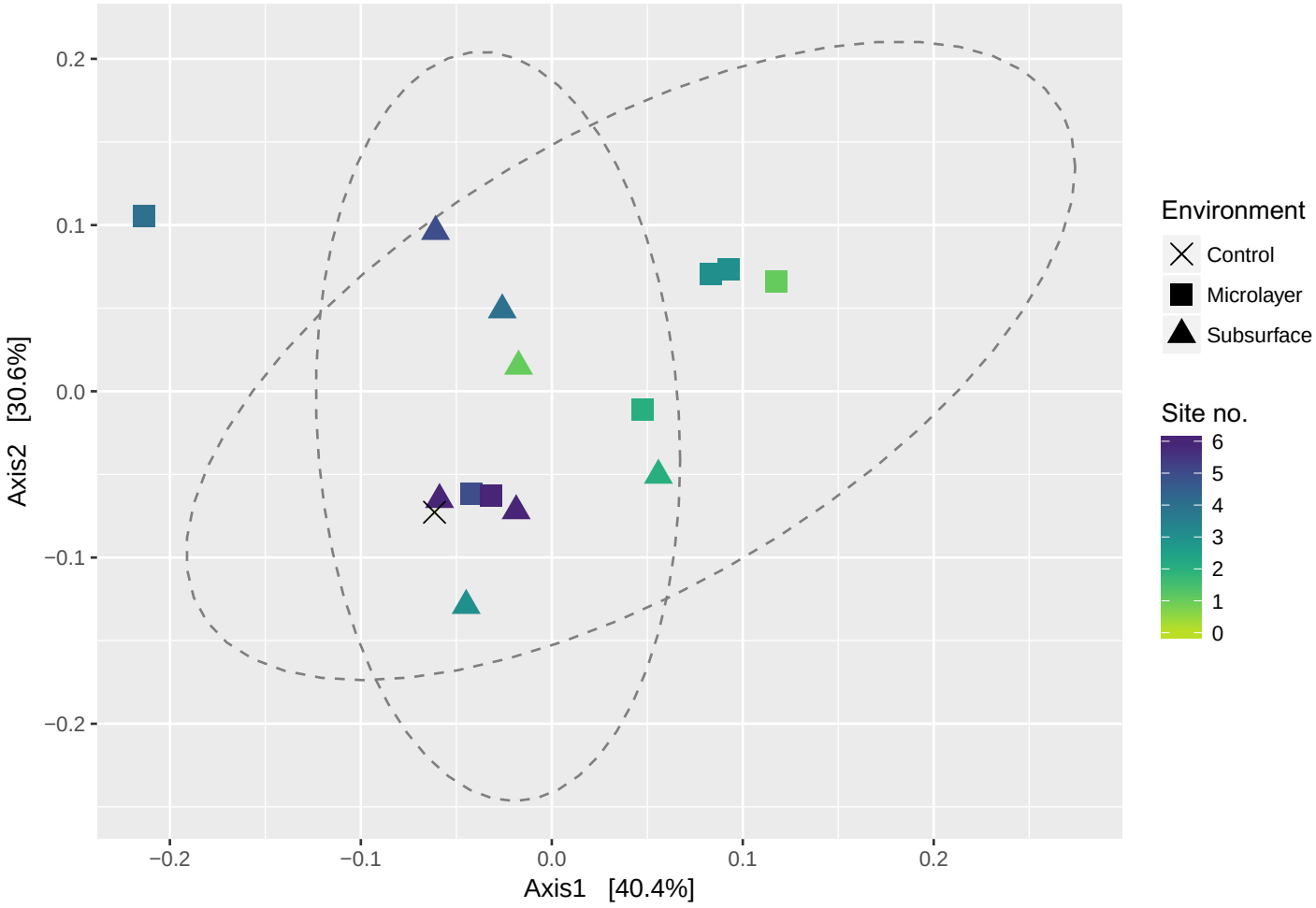

Figure 10. Double principle coordinates analysis (DPCoA) with colour-coded site identity ($n = 15$). In order to account for the first eigenvalue being always higher than the second (i.e. amount of variance explained), the axis heights were normalised to the respective eigenvalue ratios.

Table 1. Field parameters and the associated metadata for collected water samples. All of the sampling was done on 18 December 2017 during high tide.

Sample	Site no.	Coordinates	Provenance	Temp. (°C)	pH	Turbidity (NTU)	DO (mg/L)	TDS (g/L)	Salinity	Wind (k/s)	TEP conc. (ug/L)	Cell count (10^5 cells/ml)
SML-1	1	55°01.675'N 1°24.393'W	Microlayer	8.94	7.81	8.4	10.8	31.4	32.5	29.1	182.84	6.0368
SML-2	2	55°00.961'N 1°22.961'W	Microlayer	9.23	7.75	0	9.58	31.8	33.2	28.3	77.71	10.4990
SML-3	3	55°00.535'N 1°22.961'W	Microlayer	8.30	7.68	15.3	8.33	27.2	26.7	22.6	209.83	3.6776
SML-3	3	55°00.535'N 1°22.961'W	Microlayer	8.30	7.68	15.3	8.33	27.2	26.7	22.6	240.97	10.5108
SML-4	4	54°59.963'N 1°26.685'W	Microlayer	8.21	7.72	17.5	7.43	25.6	24.7	8.1	222.40	8.0181
SML-5	5	54°59.246'N 1°28.419'W	Microlayer	7.70	7.66	32.4	7.39	21.1	19.5	23.9	230.06	2.3440
SML-6	6	54°57.664'N 1°32.699'W	Microlayer	7.22	7.58	48.1	8.24	12	9.5	17.2	94.77	1.5261
SSW-1	1	55°01.675'N 1°24.393'W	Subsurface	8.94	7.81	8.4	10.8	31.4	32.5	29.1	130.75	3.0774
SSW-2	2	55°00.961'N 1°22.961'W	Subsurface	9.23	7.75	0	9.58	31.8	33.2	28.3	38.89	9.6303
SSW-3	3	55°00.535'N 1°22.961'W	Subsurface	8.30	7.68	15.3	8.33	27.2	26.7	22.6	182.45	12.7202
SSW-4	4	54°59.963'N 1°26.685'W	Subsurface	8.21	7.72	17.5	7.43	25.6	24.7	8.1	425.83	3.5968
SSW-5	5	54°59.246'N 1°28.419'W	Subsurface	7.70	7.66	32.4	7.39	21.1	19.5	23.9	235.47	3.8278
SSW-6	6	54°57.664'N 1°32.699'W	Subsurface	7.22	7.58	48.1	8.24	12	9.5	17.2	55.52	1.5967
SSW-6	6	54°57.664'N 1°32.699'W	Subsurface	7.22	7.58	48.1	8.24	12	9.5	17.2	124.89	3.2988
CTRL	0	NA	Control	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 2. Illumina 16S sequencing and preliminary data processing summaries. Unless stated otherwise, the values represent number of raw sequence reads.

Sample	Input	Filtered	Denoised	Merged	Tabled	Non-chimeric	Chimeric (%)	Total removed (%)
CTRL	29080	25672	25672	23062	23062	22965	0.42	21.03
SML-1	127987	125664	125664	124079	124079	123689	0.31	3.36
SML-2	222590	218603	218603	214589	214589	213964	0.29	3.88
SML-3	222840	219635	219635	215653	215653	213589	0.96	4.15
SML-3	77969	75344	75344	73570	73570	73163	0.55	6.16
SML-4	169991	164480	164480	157595	157595	154760	1.80	8.96
SML-5	49695	47708	47708	46801	46801	46341	0.98	6.75
SML-6	32354	29205	29205	28346	28346	28284	0.22	12.58
SSW-1	65243	61107	61107	59463	59463	59379	0.14	8.99
SSW-2	204172	199504	199504	197011	197011	196365	0.33	3.82
SSW-3	269677	264871	264871	259049	259049	247299	4.54	8.30
SSW-4	76255	73583	73583	72431	72431	72355	0.10	5.11
SSW-5	81153	77995	77995	76466	76466	76205	0.34	6.10
SSW-6	69938	65814	65814	63367	63367	63187	0.28	9.65
SSW-6	33851	31115	31115	29531	29531	29362	0.57	13.26