

# Ecological processes shaping highly connected bacterial communities along strong environmental gradients

Wenxue Wu<sup>1,2,3,\*</sup>, Chih-hao Hsieh<sup>4</sup>, Ramiro Logares<sup>5</sup>, Jay T. Lennon<sup>6</sup>, Hongbin Liu<sup>2,7,\*</sup>

<sup>1</sup>State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou 570228, Chinese mainland

<sup>2</sup>Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai 519082, Chinese mainland

<sup>3</sup>School of Marine Science, Sun Yat-sen University, Zhuhai 519082, Chinese mainland

<sup>4</sup>Institute of Oceanography, National Taiwan University, Taipei 106319, Taiwan

<sup>5</sup>Institute of Marine Sciences, CSIC, Barcelona 08003, Spain

<sup>6</sup>Department of Biology, Indiana University, Bloomington, IN 47405, United States

<sup>7</sup>Department of Ocean Science, The Hong Kong University of Science and Technology, Kowloon 999077, Hong Kong

\*Corresponding authors. Wenxue Wu, State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou 570228, Chinese mainland. E-mail: [wuwu@hainanu.edu.cn](mailto:wuwu@hainanu.edu.cn); Hongbin Liu, Department of Ocean Science, The Hong Kong University of Science and Technology, Kowloon 999077, Hong Kong. E-mail: [liuhb@ust.hk](mailto:liuhb@ust.hk)

Editor: [Julie Olson]

## Abstract

Along the river–sea continuum, microorganisms are directionally dispersed by water flow while being exposed to strong environmental gradients. To compare the two assembly mechanisms that may strongly and differently influence metacommunity dynamics, namely homogenizing dispersal and heterogeneous selection, we characterized the total (16S rRNA gene) and putatively active (16S rRNA transcript) bacterial communities in the Pearl River–South China Sea Continuum, during the wet (summer) and dry (winter) seasons using high-throughput sequencing. Moreover, well-defined sampling was conducted by including freshwater, oligohaline, mesohaline, polyhaline, and marine habitats. We found that heterogeneous selection exceeded homogenizing dispersal in both the total and active fractions of bacterial communities in two seasons. However, homogeneous selection was prevalent (the dominant except in active bacterial communities during summer), which was primarily due to the bacterial communities' tremendous diversity (associated with high rarity) and our specific sampling design. In either summer or winter seasons, homogeneous and heterogeneous selection showed higher relative importance in total and active communities, respectively, implying that the active bacteria were more responsive to environmental gradients than were the total bacteria. In summary, our findings provide insight into the assembly of bacterial communities in natural ecosystems with high spatial connectivity and environmental heterogeneity.

**Keywords:** metacommunity; heterogeneous selection; homogenizing dispersal; metabolic activity; river–sea continuum

## Introduction

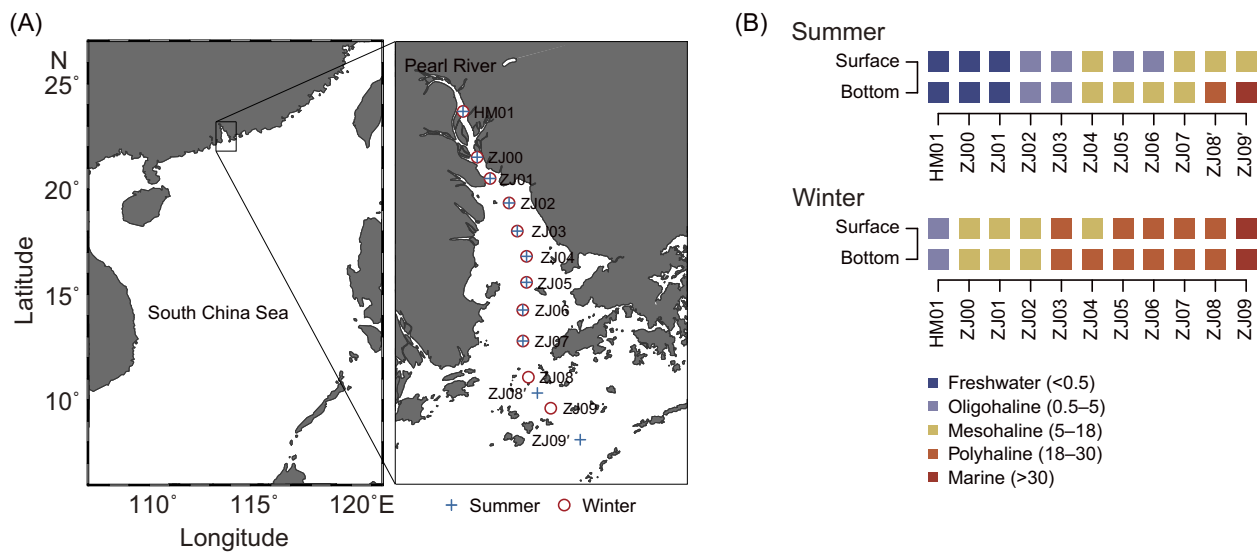
Microorganisms exhibit diverse and complex community structures across aquatic ecosystems. Although microorganisms can now be rapidly and accurately identified using high-throughput sequencing, ecological processes, which refer to the assembly of microbial communities remain complicated, especially in flowing environments where individuals are strongly influenced by currents and physical forces (Read et al. 2015, Savio et al. 2015, Henson et al. 2018, Liu et al. 2018, Gweon et al. 2021). Such conditions are common in rivers that are commonly characterized by strong environmental gradients. Therefore, the microbial communities in rivers are ideal models for quantitative metacommunity studies focusing on a set of local communities linked by dispersing individuals (Leibold et al. 2004, Vellend 2010). The metacommunity framework allows for tests of the relative importance of dispersal and selection on assembly processes (Stegen et al. 2013). Specifically, when movement is restricted, dispersal limitation leads to compositional variation among sites (Langenheder and Lindström 2019). When movement is relatively unrestricted, high rates of dispersal can homogenize compositions among sites (Heino et al. 2015). Conversely, ecological selection occurs when there are

fitness differences between organisms that either constrain (i.e. homogeneous selection) or promote (i.e. heterogeneous selection) turnover in composition among sites with respect to phylogenetic relatedness of taxa (Zhou and Ning 2017).

It is particularly challenging to understand the ecological processes that underpin the bacterial communities involved in complex biogeochemical processes at river mouths (Raymond and Bauer 2001). Many studies have shown that bacterial communities in river–sea continua change dramatically across space (Alonso et al. 2010, Jeffries et al. 2016, Aguirre et al. 2017, Sia et al. 2019). Dispersal- and selection-inferred processes may contribute inversely to bacterial community variations in river–sea continua (Stadler and del Giorgio 2022). In detail, a river–sea continuum can be defined as a reach that characteristically ranges from freshwater to seawater. Transitions typically occur within a 100–200 km distance (Wu and Liu 2018), which is a relatively small geographical scale given that bacteria can easily disperse across such highly connected ecosystems (Heino et al. 2015). As a result of high rates of immigration and emigration, dispersal may homogenize the bacterial communities (Niño-García et al. 2016, Shen et al. 2018, Fodelianakis et al. 2019). However, river–sea

Received 19 August 2024; revised 18 October 2024; accepted 29 October 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of FEMS. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** (A) Map showing the sampling locations in the Pearl River–South China Sea Continuum during the summer (indicated by blue crosses) and winter (indicated by red circles) seasons, and (B) the type of habitats (classified according to salinity shown in brackets) that each sampling location belongs to. Nine of the 11 stations sampled in each season overlap, while the remaining two are differently located. The map was generated using the Ocean Data View software (<http://odv.awi.de>).

continua are also characterized with strong environmental gradients involving variables like salinity (Langenheder et al. 2003, Paver et al. 2018), which can efficiently differentiate bacterial communities harboring taxa with distinct niches (Lozupone and Knight 2007, Thompson et al. 2017). Heterogeneous selection can emerge as a result of promoted turnover among bacteria with distantly phylogenetic relatedness, and may overpower the role of homogenizing dispersal. Although river–sea continua are an important link between terrestrial, freshwater, and marine ecosystems (Zutic and Legovic 1987, Telesh and Khlebovich 2010, Xenopoulos et al. 2017), the interaction of the underlying mechanisms that shape bacterial communities remains unclear.

The relative importance of certain ecological processes can differ based on the metabolic activity of microbial taxa (Muscarella et al. 2016, Wisnoski et al. 2020). To gain a comprehensive understanding and account for metabolic states, total and active bacterial communities are increasingly being studied together (Jones and Lennon 2010, Campbell et al. 2011, Lennon and Jones 2011, Campbell and Kirchman 2013, Richa et al. 2017, Locey et al. 2020). In some aquatic ecosystems, only a small proportion of the total bacterial community is potentially active, ranging from <5% (e.g. in oligotrophic open oceans) to >50% (e.g. in highly productive estuaries) (del Giorgio and Scarborough 1995, Smith and del Giorgio 2003). As a result, the total and active bacterial communities exhibit distinct structures. For example, the total community may contain slow-growing, dormant, and dead bacteria that can be transported across strong environmental gradients (Lennon et al. 2018, Nagler et al. 2018). Their reduced metabolic state may improve colonization success and dampen spatial variability of community structure (Carini et al. 2016), thereby reducing the expected signature of heterogeneous selection. Conversely, active communities exhibit faster and greater responses to changing environmental conditions (Hoshino and Matsumoto 2007, Franzosa et al. 2015, De Vrieze et al. 2016), and homogeneous selection that constrains community structure might be negated. Although many studies have described total and active bacterial communities concurrently, the differences in their assembly mechanisms were only compared in a small portion of previous studies (Logue

and Lindström 2010, Jia et al. 2020, Locey et al. 2020, Stadler and del Giorgio 2022).

Here, we study the bacterial communities in the Pearl River–South China Sea Continuum (PSC). According to the Venice System for the classification of marine waters (Anonymous 1958), the PSC covers diverse habitats spanning a range of salinities, including freshwater (<0.5), oligohaline (0.5–5), mesohaline (5–18), polyhaline (18–30), and marine (>30), for approximate 120 km. Moreover, the PSC is strongly influenced by seasonal river runoff events (Sun et al. 2014), which may affect the balance between dispersal- and selection-inferred processes (Huber et al. 2020). We conducted surveys along the PSC during both summer (wet season) and winter (dry season). At each sampling station, we collected surface and bottom samples and characterized bacterial communities using high-throughput sequencing of the 16S rRNA genes (DNA) and transcripts (RNA). Despite several debatable limitations and caveats (Blazewicz et al. 2013), it is generally accepted that DNA-based data represent the total bacterial communities (including living, dormant, and dead organisms), whereas RNA-based data primarily account for putatively active bacteria (Kirchman 2016). We quantified the relative importance of five ecological processes: heterogeneous selection, dispersal limitation, undominated fraction, homogenizing dispersal, and homogeneous selection, using a two-step framework (Stegen et al. 2013, Stegen et al. 2016). We compared homogenizing dispersal versus heterogeneous selection in each of the four sets of bacterial communities (i.e. total and active bacteria in summer and winter seasons), and tested how homogeneous and heterogeneous selection (in each season) varied in total versus active communities. This study provides novel insights into the assembly of highly connected bacterial communities along strong environmental gradients.

## Materials and methods

### Field sampling of bacterial communities

Water samples for molecular analyses were taken from the PSC in 2014 during both the wet summer (i.e. between May 13 and 15) and dry winter (between December 29 and 31) (Fig. 1A). In each

season, a total of 22 samples were collected from the surface (at 0.5 m depth) and bottom (5–21 m in the summer and 4.5–14 m in the winter; i.e. at 0.5–2 m above the sediment) layers across a transect (11 samples per layer). During summer and winter, 9 of the 11 stations were repeatedly surveyed, while the other two stations had a shift in geographical locations between two seasons. Considering context dependency in metacommunity studies (Heino et al. 2012), our PSC sampling design covered sequential freshwater, oligohaline, mesohaline, polyhaline, and marine habitats (according to salinity therein). After pre-filtering 100–800 ml water through an 80- $\mu$ m mesh, bacterial cells were collected using 0.2- $\mu$ m polycarbonate membranes (Millipore, Carrigtwohill, Cork, Ireland). After immersing the filters in RNA *later* solution (Ambion, Austin, TX, USA), they were immediately frozen in liquid nitrogen before being stored at  $-80^{\circ}\text{C}$  until molecular analyses.

### Environmental distance calculation

To determine environmental contexts across the transect, we included a series of abiotic and biotic environmental variables, including temperature, salinity, chlorophyll *a*,  $\text{NH}_4$ ,  $\text{NO}_2$ ,  $\text{NO}_3$ ,  $\text{PO}_4$ , suspended particulate matter, particulate organic carbon, particulate nitrogen, and cell numbers of heterotrophic bacteria, *Synechococcus*, and pigmented picoeukaryotes, which were analyzed as described previously (Wu and Liu 2018). Pearson's correlations between environmental variables were analyzed and visualized via heatmaps.

To determine environmental distances between sampling sites, the values of each environmental variable were standardized with a mean of 0 and variance of 1, and the environmental distances were then calculated as the Euclidean distance between sites. This transformation enabled the linking of ecological processes (estimated as below) with environmental contexts (and hence, to a holistic environmental perspective) rather than to a single variable. The relationship between environmental and geographical distances was evaluated using Mantel test.

### Library construction and sequence processing

DNA/RNA extraction and cDNA synthesis procedures have been previously described in detail (Wu and Liu 2018). Barcoded primers 515F (5'-GTGCCAGCMGCCGCGTAA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') (Caporaso et al. 2011) were used to amplify the V4 region of the 16S rRNA (gene), based on a dual-index strategy. Polymerase chain reaction (PCR) mixtures of 20  $\mu$ l contained 1 $\times$  PCR buffer, 1.5 mM  $\text{MgCl}_2$ , 0.2 mM dNTP mix, 0.5  $\mu$ M each of primer, 1 U Invitrogen Platinum *Taq* DNA polymerase (Life Technologies, Carlsbad, CA, USA), and DNA/cDNA templates ( $\sim$ 1–10 ng). PCR was performed as follows:  $94^{\circ}\text{C}$  for 3 min, followed by 30 cycles at  $94^{\circ}\text{C}$  for 45 s,  $50^{\circ}\text{C}$  for 60 s, and  $72^{\circ}\text{C}$  for 90 s, with a final extension step at  $72^{\circ}\text{C}$  for 10 min. For each sample, triplicate PCR products were pooled and sequenced on a HiSeq 2500 platform (Illumina, San Diego, CA, USA), generating 2  $\times$  250 bp paired-end reads.

The Quantitative Insights Into Microbial Ecology (QIIME v. 1.9.1) pipeline was used to process the sequences (Caporaso et al. 2010). Using the default parameters, overlapping paired-end reads were merged using SeqPrep (<https://github.com/jstjohn/SeqPrep>). The merged sequences were then filtered using the following criteria: 1) Phred scores < 25; 2) homopolymers > 6; 3) with any mismatch in the primers; 4) with any error in the barcodes; 5) containing ambiguous bases; 6) too short (<270 bp) or too long (>320 bp). The filtered sequences were grouped into operational taxonomic units (OTUs) using Sumacust (Mercier et al. 2013) at 99% simi-

larity (Koeppel and Wu 2014, Louca et al. 2016). After removing singletons and doubletons, ChimeraSlayer was used to identify and remove chimeras based on representative sequences (i.e. the most abundant sequence in each OTU) (Haas et al. 2011). These representative sequences were annotated against the SILVA 119 database (Quast et al. 2013) using BLAST ( $E$  value =  $10^{-6}$ ), and those affiliated with archaea, eukaryotes, chloroplasts, and mitochondria were removed from the OTU table. The remaining representative sequences were aligned with MAFFT 7 (Katoh and Standley 2013) using the FFT-large-NS-2 method. After the removal of columns that showed gaps in >90% of the positions, the resulting alignment was used to construct a phylogenetic tree with FastTree (Price et al. 2009).

For further analysis, the bacterial communities (full OTU tables) were normalized to the minimum sequencing depth (i.e. 5557 sequences) associated with 100 bootstrap resampling runs. The number of observed OTUs (OTU richness) was calculated using the vegan package (Oksanen et al. 2018) in R v. 3.5.0 (R Core Team 2018). The relationship between OTU richness (based on an average of 100 bootstraps) and distance to the uppermost end (i.e. Station HM01) was tested using linear regression or generalized additive model in R. The rarefied bacterial communities were also used to generate the weighted UniFrac dissimilarity, which was further plotted using principal coordinate analysis (PCoA) using the vegan package.

### Quantitative estimate of assembly processes

To quantitatively examine dispersal- and selection-inferred ecological processes, we used the null model approach (Stegen et al. 2016) separately for the four sets (total and active; summer and winter) of bacterial communities. First, we tested the phylogenetic signals using Mantel correlograms to determine the relationship between phylogenetic distance and niche differences (inferred from all environmental variables) among OTUs prior to conducting null model analyses (Stegen et al. 2013, Dini-Andreote et al. 2015). Only OTUs (in the full OTU tables) with an occupancy (i.e. number of sites occupied) of >3 were included when testing the phylogenetic signals because an extremely low occupancy might bias the detection of niche specialization. Then, the phylogenetic turnover in the (rarefied) bacterial communities was quantified based on  $\beta$ -mean nearest taxon distance ( $\beta$ MNTD) using the picante package (Kembel et al. 2010). To infer the relative importance of different ecological processes, we calculated the pairwise  $\beta$ -nearest taxon index ( $\beta$ NTI), which accounted for the difference between the observed  $\beta$ MNTD and the mean of the null distribution normalized by its standard deviation (999 randomizations).  $\beta$ NTI values of  $\leq -2$  and  $\geq +2$  indicate significantly less than and more than the expected phylogenetic turnover, respectively (i.e. homogeneous and heterogeneous selection, respectively). In cases where  $|\beta$ NTI| < 2, we calculated the Raup–Crick metric based on Bray–Curtis dissimilarities ( $\text{RC}_{\text{bray}}$ ) (Stegen et al. 2013). Pairwise  $\text{RC}_{\text{bray}}$  values of >+0.95, <−0.95, and  $|\text{RC}_{\text{bray}}| < 0.95$  accounted for the relative contributions of dispersal limitation, homogenizing dispersal, and the undominated fraction (mainly indicating drift under weak dispersal- and selection-inferred processes), respectively (Dini-Andreote et al. 2015, Stegen et al. 2015, Zhou and Ning 2017). The bootstrap analysis for 100 replicates was included in the null model analyses.

As described above, when determining a driving ecological process for pairwise bacterial communities, the number of OTUs (corresponding to the nodes of the phylogenetic tree used for the calculation of  $\beta$ MNTD) and their relative abundances are

important. This is because these two community characteristics can essentially affect the values of  $\beta$ NTI and  $RC_{\text{bray}}$  based on 999 randomizations. To test the effect of OTU numbers and relative abundances, we further calculated ecological processes shaping the sub-communities composed of the most abundant 1000 OTUs. The resulting difference in ecological processes (compared to those of whole bacterial communities) supports the effect of diversity (i.e. OTU numbers) and rarity (i.e. relative abundances of OTUs) on null model analyses.

### Habitat preference of OTUs

To identify OTUs that are specifically associated with different habitats (i.e. freshwater, oligohaline, mesohaline, polyhaline, and marine environments), the analysis of indicator value (IndVal) was performed based on relative abundances (in the full OTU tables) using the *indicpecies* package (De Cáceres and Legendre 2009). The OTUs with a preference of a particular habitat were recognized as having significant ( $P < 0.05$ ) IndVal estimations (ranging from 0 to 1) through 999 permutations (Hauptmann et al. 2016). OTUs showing a relative abundance of  $<0.01\%$  in all the samples were removed before IndVal analyses because their rarities biased the prediction of habitat preference.

### Statistical analysis

To visualize community characteristics, rank–abundance and rarefaction curves were plotted using the *vegan* package with the *radfit* and *rarefy* functions, respectively. Moreover, the importance of each environmental factor in promoting beta diversity (weighted UniFrac dissimilarity) was individually tested using permutational multivariate analysis of variance (PERMANOVA, 999 permutations) with the *adonis2* function in the *vegan* package. The PERMANOVA was also used to examine the difference between the total and active bacterial communities (999 permutations). To assess the community assembly processes along the PSC, we regressed pairwise  $\beta$ NTI values (using an average of 100 bootstraps) against environmental distances (Kumari et al. 2015, Stegen et al. 2016, Feng et al. 2018, Zhang et al. 2018).

## Results

### Hydrographic condition

There were a total of 49 and 37 significant correlation relationships of the 13 environmental variables in summer and winter, respectively (Supplementary Fig. S1). As a typical hydrographic context in this river–sea continuum, salinity showed significantly negative correlations with nutrients ( $\text{NH}_4$ ,  $\text{NO}_2$ ,  $\text{NO}_3$ , and  $\text{PO}_4$ ), particulate organic carbon, and particulate nitrogen in both seasons. Environmental distances among sites significantly increased with geographical distances based on Mantel tests (summer:  $r = 0.77$ ,  $P < 0.001$ ; winter:  $r = 0.76$ ,  $P < 0.001$ ). Moreover, a total of 6 and 0 samples located in the freshwater habitat in summer and winter, respectively (Fig. 1B), indicating that this transect was more influenced by freshwater in the wet season due to more river runoffs.

### Community characteristic

The sum of OTUs detected in summer and winter were 45 594 and 40 328, respectively, a considerable percentage of which was shared (summer, 43%; winter, 49%) between total and active bacterial communities (Supplementary Fig. S2A and B). The rank–abundance curves displayed a long tail (Supplementary Fig. S2C and D), and the rarefaction curves of individual bacterial commu-

nities generally did not reach an asymptote (Supplementary Fig. S2E).

Total bacterial communities showed significantly ( $P < 0.05$ ) higher OTU richness than their active counterparts (Supplementary Fig. S3). Moreover, OTUs in total bacterial communities had higher habitat occupancies than those in active fractions, when compared based on a given relative abundance (Supplementary Fig. S4). Remarkably, the most abundant 10 OTUs in each dataset, except an OTU in active bacterial communities in summer (OTU49), were detected in all samples (Supplementary Fig. S5).

The PCoA ordination revealed that total and active bacterial communities were significantly different ( $P < 0.001$ ), both of which shifted along salinity gradients (Fig. 2). Furthermore, salinity was the only environmental factor that was significantly ( $P < 0.05$ ) related to community dissimilarities in all four sets of bacterial communities (Supplementary Table S1).

### Community assembly

The null model analyses, supported by the phylogenetic signal particularly pronounced among OTUs with short phylogenetic distances (Supplementary Fig. S6), showed that ecological processes on average comprised 38.2% homogeneous selection, 8.5% homogenizing dispersal, 21.8% undominated fraction (indicating weak dispersal- and selection-inferred processes), 15% dispersal limitation, and 16.5% heterogeneous selection (Fig. 3). In detail, heterogeneous selection exceeded homogenizing dispersal in all four sets of bacterial communities, in spite of a slightly different proportion in the active bacterial communities in winter (heterogeneous selection, 11.3%; homogenizing dispersal, 10%).

In either summer or winter seasons, homogeneous and heterogeneous selection showed higher proportions in total and active bacterial communities, respectively (Fig. 3). Notably, ecological processes shifted with changes in environmental distance (Fig. 4), characterized by significant linear regression relationships ( $P < 0.001$ ) between environmental distances and  $\beta$ NTI values (Supplementary Fig. S7).  $\beta$ MNTD also increased with increasing environmental distances (Supplementary Fig. S8).

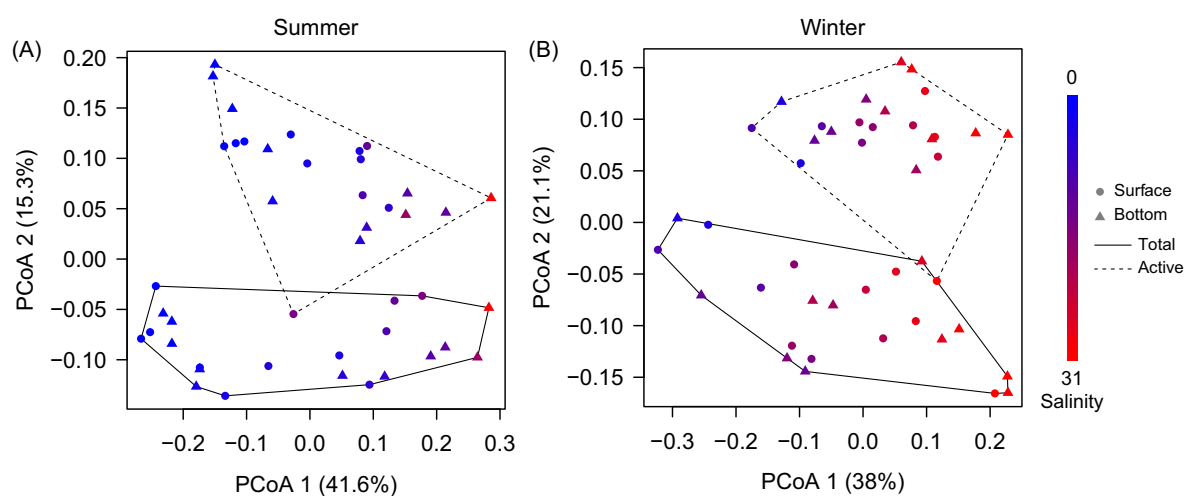
The assembly of the sub-communities of the most abundant 1000 OTUs was on average comprised of 0.04% homogeneous selection, 9.8% homogenizing dispersal, 23.8% undominated fraction, 21.7% dispersal limitation, and 44.6% heterogeneous selection (Supplementary Fig. S9). Moreover, heterogeneous selection was the highest except in the total bacterial communities in winter showing an undominated fraction of 39.3% (heterogeneous selection = 31.5%).

In addition, accompanied with the difference in ecological processes between total and active bacterial communities (Fig. 3), the former showed less OTUs with habitat preference than the latter in both summer (total, 63; active, 323) and winter (total, 149; active, 178) (Fig. 5).

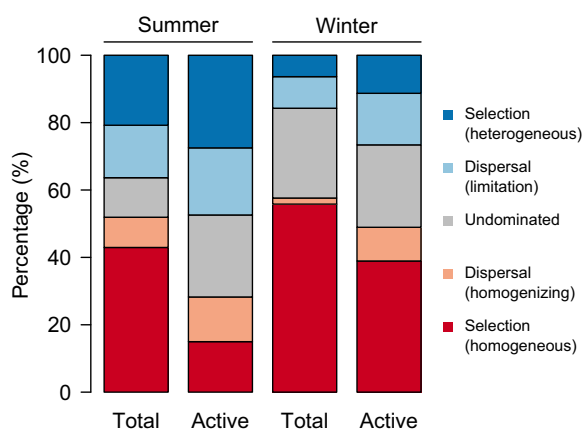
## Discussion

Our study found that heterogeneous selection exceeded homogenizing dispersal in shaping bacterial communities in the PSC, implying that environmental gradients are more important than geographic connections. In each season, homogeneous and heterogeneous selection showed higher contributions in total and active bacterial communities, respectively, supporting that the metabolic activity makes a difference in bacterial community assembly.





**Figure 2.** PCoA plots based on weighted UniFrac dissimilarities during (A) summer and (B) winter associated with changes in salinity, which are indicated by colors such that blue is 0 and red is 31. The total and putatively active bacterial communities found in the surface and bottom samples collected at each season are indicated by different lines and symbols, respectively. Each point represents the average derived from 100 bootstraps. The only environmental factor that was significantly ( $P < 0.05$ ) related to community dissimilarities in all four sets of bacterial communities was salinity (Supplementary Table S1). The total and active bacterial communities were shown to be significantly different using PERMANOVA (summer,  $R^2 = 0.14$ ,  $P < 0.001$ ; winter,  $R^2 = 0.2$ ,  $P < 0.001$ ).



**Figure 3.** Barplots showing the proportions of ecological processes (based on 100 bootstraps) in shaping total and active bacterial communities in summer and winter seasons.

### Ecological process shifting in the PSC

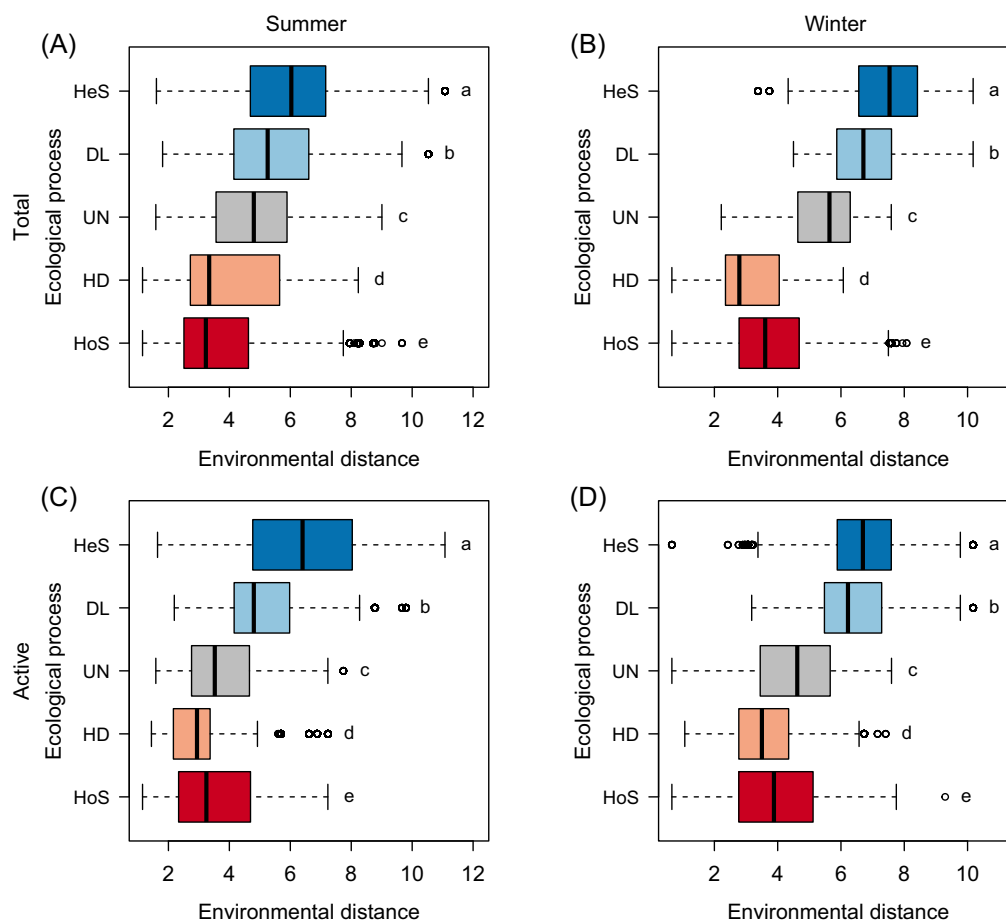
Ecological processes shaping bacterial communities mainly occurred under different environmental distances in the PSC (Fig. 4 and Supplementary Fig. S7). We divide the environmental distances into three levels for ease of interpretation: small (mainly showing homogeneous selection and homogenizing dispersal), moderate (mainly showing undominated fraction), and large (mainly showing dispersal limitation and heterogeneous selection). Specifically, under small environmental distances (accompanied by small geographical distances), the passive dispersal of bacteria can be efficiently facilitated (Heino et al. 2015). However, homogenizing dispersal has to compete against homogeneous selection in this scenario (Souffreau et al. 2014). Homogeneous selection was prevalent (Fig. 3) based on estimations in the first step of the null model analysis, which therefore weakens the role of homogenizing dispersal determined in the second step.

In contrast, the large environmental distances act as strong selective forces for bacterial taxa with extremely distinct evolutionary histories showing scarce freshwater-marine transitions (Loga-

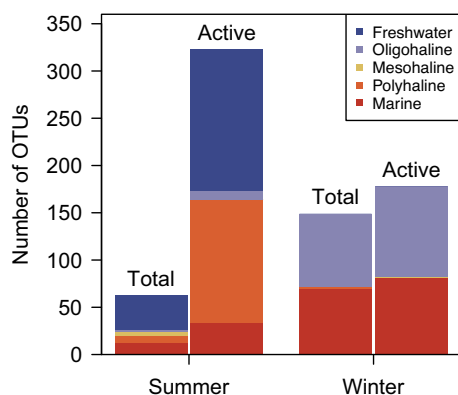
res et al. 2009, Newton et al. 2011, Herlemann et al. 2016, Paver et al. 2018). For example, selective pressures such as salinity-related mortality (Painchaud et al. 1995) can efficiently differentiate bacterial communities (Fig. 2). Moreover, distinct trends arose at the OTU level in freshwater- and seawater-like habitats with relative abundances of the most abundant OTUs peaking at different locations across the transect (Supplementary Fig. S5). These differential patterns under large environmental distances can be partly attributable to heterogeneous selection primarily responsible for the large phylogenetic turnover between bacterial communities (Supplementary Fig. S8). Alternatively, dispersal limitation emerged (Fig. 4 and Supplementary Fig. S7) because the majority of bacterial individuals are unable to disperse across strong environmental gradients. Therefore, our results support that dispersal limitation can arise not only from geographical barriers but also through interactions with environmental heterogeneities (Adler et al. 2007).

The undominated fraction was common among sites with moderate environmental distances (Fig. 4 and Supplementary Fig. S7). On the one hand, sites with moderate environmental distances are linked via water flow, which can greatly reduce dispersal limitation but may not be strong enough to generate homogenizing dispersal. On the other hand, selective forces such as changing salinity (Lozupone and Knight 2007) under moderate environmental distances can largely reduce homogeneous selection, but may not be strong enough to generate heterogeneous selection. Most bacteria from freshwater are stressed when transported to downstream stations, and may appear to be randomly filtered in rapidly changing environments (i.e. by drift). Therefore, the occurrence of undominated fraction is responsible for the scenario in which no single process inferred from either dispersal or selection is dominant (Zhou and Ning 2017).

In our bootstrapped results ( $n = 100$ ), the assembly of the same community pair (i.e. dots sharing an x-axis, Supplementary Fig. S7) could be attributed to different ecological processes (i.e. in different colors). To some degree, this is consistent with the idea of reconciling dispersal- and selection-based perspectives in metacommunity ecology (Gravel et al. 2006); that is, dispersal and selection can be jointly responsible for differences between each



**Figure 4.** Boxplots showing the environmental distances in which each ecological process was observed in the (A, B) total and (C, D) active bacterial communities. The boxplots display five summary statistics, including median (black line), 25% and 75% percentiles (box), and range (whiskers), and outliers are indicated by circles. The different letters (a–e) next to each boxplot indicate the significant differences in pairwise comparisons based on the Kruskal–Wallis rank sum test ( $P < 0.001$ ). The ecological process showing a mean in the middle within the groups of environmental distances, are simply designated as “occurring at moderate environmental distances.” Accordingly, ecological processes showing a higher or lower mean are defined as “occurring at high or low environmental distances,” respectively. HeS, heterogeneous selection; DL, dispersal limitation; UN, undominated fraction; HD, homogenizing dispersal; and HoS, homogeneous selection.



**Figure 5.** Barplots showing the numbers of OTUs with habitat preference in the five types of habitats (freshwater, oligohaline, mesohaline, polyhaline, and marine).

pair of bacterial communities, despite only one ecological process recognized in null model analyses. Alternatively, rarefaction is a random process that can cause instabilities in null model analyses (Ma and Tu 2022). That is, rarefying the largest sample of 59 963 sequences into 5557 sequences (i.e. minimum sequencing

depth) may result in some compositional instability among the subsampled communities, influencing the ecological processes in the subsequent analyses. Nonetheless, we suggest that including a bootstrap step with repeatedly subsampling (Schloss 2024) can get more accurate quantifications of ecological processes.

### Prevalence of homogeneous selection

Our results showed that homogeneous selection can emerge in ecosystems with strong environmental gradients, consistent with several recent studies with similar observations (Wang et al. 2020, Wu et al. 2021, Urvoy et al. 2022, Blais et al. 2024). Moreover, our study expands the prevalence of homogeneous selection that has been observed in the oligotrophic South Pacific Gyre (Allen et al. 2020), hydrologically connected rivers (Graham et al. 2017, Cai et al. 2019), glacier-fed streams (Fodelianakis et al. 2022), and proglacial floodplain streams (Brandani et al. 2023).

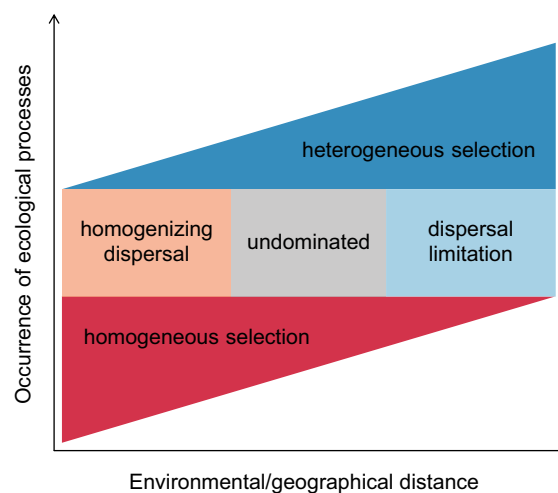
A few characteristics related to bacterial communities can primarily enhance the strength of homogeneous selection. First, the bacterial communities in the PSC are tremendously diverse in part because they received allochthonous inputs of microorganisms from soil and sediments, as well as a mix of freshwater and marine taxa (Crump et al. 2012). The PSC had a large number of OTUs during summer (total, 41 582; active, 23 611) and winter

(total, 34 309; active, 25 790). Meanwhile, both total and active bacterial communities exhibited a high degree of rarity (indicated by the long tails in the rank–abundance curves, [Supplementary Fig. S2C and D](#)), which is a common property in bacterial communities (Pedrós-Alió 2012). Second, sites with small environmental distances may share some OTUs with high relative abundances (for convenience, simply designated as abundant OTUs). Taking the most abundant 10 OTUs for example, they commonly showed comparable relative abundances between neighboring sites ([Supplementary Fig. S5](#)). As a result, the phylogenetic distances between two bacterial communities (quantified by the observed  $\beta$ MNTD) under small environmental distances were relatively low ([Supplementary Fig. S8](#)) because this metric is determined by the phylogenetic positions and relative abundances of OTUs. More importantly, the null expectation for pairwise community distances is derived by randomizing OTUs across the phylogeny and recalculating  $\beta$ MNTD 999 times (Stegen et al. 2012), and thus the observed  $\beta$ MNTD values may be much lower than the null expectation based on tremendous diversity associated with high rarity. A resulting high degree of deviation of  $\beta$ MNTD from the null expectation (i.e.  $\beta$ NTI  $\leq -2$ ) is a sign of homogeneous selection.

When focusing on sub-communities composed of the most abundant 1000 OTUs (i.e. lowering the diversity and rarity), we found that homogeneous selection was significantly negated with contributions approximately equal to 0 ([Supplementary Fig. S9](#)). Similar to our examination regarding diversity associated with rarity, a recent study conducted in inshore waters showed the dominance of non-selection processes in shaping bacterial sub-communities of the most abundant 500 OTUs (Han et al. 2022). This is because both diversity and rarity are intensively lowered in null modeling, leading to a low degree of deviation of  $\beta$ MNTD from the null expectation (i.e.  $|\beta$ NTI|  $< 2$ ). Another recent study showed that the relative importance of homogeneous selection increases with increasing similarity cutoffs for grouping OTUs (Quiroga et al. 2022). This is primarily due to the fact that using a higher similarity cutoff (i.e. a higher phylogenetic resolution) can largely increase both diversity (i.e. having an expanded phylogenetic tree when calculating  $\beta$ MNTD) and rarity (i.e. lowering the sequence number per OTU) for a given dataset. In addition, the sharing of abundant OTUs between sites (as mentioned above) benefits from our gradient-type sampling design (including freshwater, oligohaline, mesohaline, polyhaline, and marine habitats). Given a coarser sampling, e.g. by only sampling the two ends of the PSC (i.e. HM01 and ZJ09), the selection would be mainly heterogeneous rather than homogeneous ([Supplementary Fig. S7](#)). This is consistent with the notion that community assembly is scale-dependent (Chase 2014, Ladau and Eloe-Fadrosh 2019, Langenheder and Lindström 2019). Therefore, we suggest that the prevalence of homogeneous selection in this study is primarily attributable to the bacterial communities' tremendous diversity (associated with high rarity) and our specific sampling design.

### Total versus active bacterial communities

As expected, homogeneous and heterogeneous selection in each season were more important in total and active bacterial communities, respectively (Fig. 3), consistent with that active bacteria are more responsive to environmental gradients than are the total bacteria (De Vrieze et al. 2016). This is supported by the observation that active bacterial communities had more OTUs showing habitat preferences than total bacterial communities (Fig. 5). Moreover, OTUs in the total communities exhibited con-



**Figure 6.** A conceptual model indicates the environmental dependency of ecological processes in a river–sea continuum. In the event of low environmental distances (commonly showing significant correlations to geographical distances), homogeneous selection and homogenizing dispersal occur. Under moderate environmental distances, neither dispersal- nor selection-inferred processes are dominant and so the fraction is undominated. Under large environmental distances, dispersal limitation and heterogeneous selection become the major processes. This progressive succession in ecological processes could result in a significant linear regression relationship between  $\beta$ NTI values and environmental distances.

sistently higher occupancies than those in the active fractions ([Supplementary Fig. S4](#)), also implying that the former could be less sensitive to environmental gradients by inhabiting a wider variety of niches. Notably, the higher occupancies of OTUs in total bacterial communities may primarily result from the inclusion of dead members, because dead bacteria can potentially persist in habitats ranging from freshwater to marine. This interpretation is in line with a study demonstrating that marine bacteria can even be recruited from freshwater sources in DNA-inferred communities (Comte et al. 2014).

However, when interpreting the differences in community assembly, it should be noted that active bacterial communities are not a simple subset of total communities. For example, a significant proportion of OTUs was found only in the active communities (summer = 8.8%, winter = 14.9%) ([Supplementary Fig. S2A and B](#)), and seven active communities had higher OTU richness than their total counterparts ([Supplementary Fig. S3](#)). The OTUs found exclusively in active communities could be extremely rare taxa from the rare biosphere (Sogin et al. 2006). The detection limitation in sequencing methodology contributes to this discrepancy because DNA-based data might lose some rare taxa that were solely recovered in RNA-based data (and vice versa). Alternatively, these OTUs could have been derived from RNA that survives outside the organismal environment (Cristescu 2019). Moreover, examining the assembly of either total or active bacterial communities is complicated and influenced by a variety of factors. For example, it is well known that both rDNA (Stoddard et al. 2015) and rRNA (Kirchman 2016) copy numbers vary greatly among bacterial taxa, with rRNA copy numbers in particular varying by several orders of magnitude for marine bacterial strains at various metabolic stages (Fegatella et al. 1998). As a result, when comparing the assembly of total and active bacterial communities, their skewed community structures are a significant source of biased interpretations (Lavrinenko et al. 2021).

## Conclusions

Our study makes inferences about how highly connected bacterial communities are assembled along strong environmental gradients (Fig. 6). Specifically, homogeneous selection and homogenizing dispersal occur primarily in small environmental distances; the undominated fraction mainly emerges in moderate environmental distances and mostly indicates drift; dispersal limitation and heterogeneous selection occur primarily in large environmental distances. In the PSC, heterogeneous selection exceeds homogenizing dispersal in all four sets of bacterial communities, and homogeneous and heterogeneous selection (in each season) are more important in total and putatively active bacterial communities, respectively. Moreover, our study reveals that calculating ecological processes can be heavily influenced by a few community characteristics (e.g. diversity and rarity). This suggests that when interpreting the ecological processes, the phylogenetic and proportional properties of a metacommunity should be carefully considered. In summary, this study emphasizes the importance of disentangling community assembly mechanisms to understand microbial biogeography in natural ecosystems.

## Author contributions

Wenxue Wu (Conceptualization, Formal analysis, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing), Chih-hao Hsieh (Formal analysis, Writing – original draft, Writing – review & editing), Ramiro Logares (Writing – original draft, Writing – review & editing), Jay T. Lennon (Writing – original draft, Writing – review & editing), and Hongbin Liu (Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing)

## Supplementary data

Supplementary data is available at [FEMSEC Journal](#) online.

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

## Funding

This study was supported by the National Natural Science Foundation of China (No. 42176143), the Natural Science Foundation of Hainan Province (No. 424MS025), the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (No. 311022005), the Area of Excellence scheme of the Hong Kong Research Grants Council (No. AoE/P-601/23-N), and the Scientific Research Foundation of Hainan University (No. RZ2300002681).

## Data availability

Raw sequencing data are available from the NCBI Sequence Read Archive under PRJNA805376.

## References

- Adler PB, HilleRisLambers J, Levine JM. A niche for neutrality. *Ecol Lett* 2007;**10**:95–104.
- Aguirre M, Abad D, Albaina A et al. Unraveling the environmental and anthropogenic drivers of bacterial community changes in the Estuary of Bilbao and its tributaries. *PLoS One* 2017;**12**:e0178755.
- Allen R, Hoffmann LJ, Larcombe MJ et al. Homogeneous environmental selection dominates microbial community assembly in the oligotrophic South Pacific Gyre. *Mol Ecol* 2020;**29**:4680–91.
- Alonso C, Gómez-Pereira P, Ramette A et al. Multilevel analysis of the bacterial diversity along the environmental gradient Río de la Plata–South Atlantic Ocean. *Aquat Microb Ecol* 2010;**61**:57–72.
- Anonymous. The Venice system for the classification of marine waters according to salinity. *Limnol Oceanogr* 1958;**3**:346–7.
- Blais M, Matveev A, Lovejoy C et al. Size-dependent community patterns differ between microbial eukaryotes and bacteria in a permafrost lake–river–sea continuum. *Limnol Oceanogr* 2024;**69**:667–80.
- Blazewicz SJ, Barnard RL, Daly RA et al. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J* 2013;**7**:2061–8.
- Brandani J, Peter H, Fodelianakis S et al. Homogeneous environmental selection structures the bacterial communities of benthic biofilms in proglacial floodplain streams. *Appl Environ Microbiol* 2023;**89**:e0201022.
- Cai W, Li Y, Shen Y et al. Vertical distribution and assemblages of microbial communities and their potential effects on sulfur metabolism in a black-odor urban river. *J Environ Manage* 2019;**235**:368–76.
- Campbell BJ, Kirchman DL. Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. *ISME J* 2013;**7**:210–20.
- Campbell BJ, Yu L, Heidelberg JF et al. Activity of abundant and rare bacteria in a coastal ocean. *Proc Natl Acad Sci USA* 2011;**108**:12776–81.
- Caporaso JG, Kuczynski J, Stombaugh J et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;**7**:335–6.
- Caporaso JG, Lauber CL, Walters WA et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 2011;**108**:4516–22.
- Carini P, Marsden PJ, Leff JW et al. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat Microbiol* 2016;**2**:16242.
- Chase JM. Spatial scale resolves the niche versus neutral theory debate. *J Veg Sci* 2014;**25**:319–22.
- Comte J, Lindström ES, Eiler A et al. Can marine bacteria be recruited from freshwater sources and the air? *ISME J* 2014;**8**:2423–30.
- Cristescu ME. Can environmental RNA revolutionize biodiversity science? *Trends Ecol Evol* 2019;**34**:694–7.
- Crumph BC, Amaral-Zettler LA, Kling GW. Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J* 2012;**6**:1629–39.
- De Cáceres M, Legendre P. Associations between species and groups of sites: indices and statistical inference. *Ecology* 2009;**90**:3566–74.
- De Vrieze J, Regueiro L, Props R et al. Presence does not imply activity: DNA and RNA patterns differ in response to salt perturbation in anaerobic digestion. *Biotechnol Biofuels* 2016;**9**:244.
- del Giorgio PA, Scarborough G. Increase in the proportion of metabolically active bacteria along gradients of enrichment in freshwater and marine plankton: implications for estimates of bacterial growth and production rates. *J Plankton Res* 1995;**17**:1905–24.
- Dini-Andreote F, Stegen JC, van Elsas JD et al. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc Natl Acad Sci USA* 2015;**112**:E1326–32.
- Fegatella F, Lim J, Kjelleberg S et al. Implications of rRNA operon copy number and ribosome content in the marine oligotrophic ultramicrobacterium *Sphingomonas* sp. strain RB2256. *Appl Environ Microbiol* 1998;**64**:4433–8.
- Feng Y, Chen R, Stegen JC et al. Two key features influencing community assembly processes at regional scale: initial state and



- degree of change in environmental conditions. *Mol Ecol* 2018;**27**:5238–51.
- Fodelianakis S, Lorz A, Valenzuela-Cuevas A et al. Dispersal homogenizes communities via immigration even at low rates in a simplified synthetic bacterial metacommunity. *Nat Commun* 2019;**10**:1314.
- Fodelianakis S, Washburne AD, Bourquin M et al. Microdiversity characterizes prevalent phylogenetic clades in the glacier-fed stream microbiome. *ISME J* 2022;**16**:666–75.
- Franzosa EA, Hsu T, Sirota-Madi A et al. Sequencing and beyond: integrating molecular “omics” for microbial community profiling. *Nat Rev Microbiol* 2015;**13**:360–72.
- Graham EB, Crump AR, Resch CT et al. Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. *Environ Microbiol* 2017;**19**:1552–67.
- Gravel D, Canham CD, Beaudet M et al. Reconciling niche and neutrality: the continuum hypothesis. *Ecol Lett* 2006;**9**:399–409.
- Gweon HS, Bowes MJ, Moorhouse HL et al. Contrasting community assembly processes structure lotic bacteria metacommunities along the river continuum. *Environ Microbiol* 2021;**23**:484–98.
- Haas BJ, Gevers D, Earl AM et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 2011;**21**:494–504.
- Han D, Shin H, Lee J-H et al. Phylogenetic diversity and spatiotemporal dynamics of bacterial and microeukaryotic plankton communities in Gwangyang Bay of the Korean Peninsula. *Sci Rep* 2022;**12**:2980.
- Hauptmann AL, Markussen TN, Stibal M et al. Upstream freshwater and terrestrial sources are differentially reflected in the bacterial community structure along a small arctic river and its estuary. *Front Microbiol* 2016;**7**:1474.
- Heino J, Grönroos M, Soininen J et al. Context dependency and metacommunity structuring in boreal headwater streams. *Oikos* 2012;**121**:537–44.
- Heino J, Melo AS, Siqueira T et al. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshw Biol* 2015;**60**:845–69.
- Henson MW, Hanssen J, Spooner G et al. Nutrient dynamics and stream order influence microbial community patterns along a 2914 kilometer transect of the Mississippi River. *Limnol Oceanogr* 2018;**63**:1837–55.
- Herlemann DPR, Lundin D, Andersson AF et al. Phylogenetic signals of salinity and season in bacterial community composition across the salinity gradient of the Baltic Sea. *Front Microbiol* 2016;**7**:1883.
- Hoshino YT, Matsumoto N. DNA- versus RNA-based denaturing gradient gel electrophoresis profiles of a bacterial community during replenishment after soil fumigation. *Soil Biol Biochem* 2007;**39**:434–44.
- Huber P, Metz S, Unrein F et al. Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. *ISME J* 2020;**14**:2951–66.
- Jeffries TC, Schmitz Fontes ML, Harrison DP et al. Bacterioplankton dynamics within a large anthropogenically impacted urban estuary. *Front Microbiol* 2016;**6**:1438.
- Jia X, Dini-Andreote F, Falcão Salles J. Comparing the influence of assembly processes governing bacterial community succession based on DNA and RNA data. *Microorganisms* 2020;**8**:798.
- Jones SE, Lennon JT. Dormancy contributes to the maintenance of microbial diversity. *Proc Natl Acad Sci USA* 2010;**107**:5881–6.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80.
- Kembel SW, Cowan PD, Helmus MR et al. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 2010;**26**:1463–4.
- Kirchman DL. Growth rates of microbes in the oceans. *Annu Rev Mar Sci* 2016;**8**:285–309.
- Koeppel AF, Wu M. Species matter: the role of competition in the assembly of congeneric bacteria. *ISME J* 2014;**8**:531–40.
- Kumari P, Choi HL, Sudiarto SIA. Assessment of bacterial community assembly patterns and processes in pig manure slurry. *PLoS One* 2015;**10**:e0139437.
- Ladau J, Elie-Fadrosh EA. Spatial, temporal, and phylogenetic scales of microbial ecology. *Trends Microbiol* 2019;**27**:662–9.
- Langenheder S, Kisand V, Wikner J et al. Salinity as a structuring factor for the composition and performance of bacterioplankton degrading riverine DOC. *FEMS Microbiol Ecol* 2003;**45**:189–202.
- Langenheder S, Lindström ES. Factors influencing aquatic and terrestrial bacterial community assembly. *Environ Microbiol Rep* 2019;**11**:306–15.
- Lavrinienko A, Jernfors T, Koskimäki JJ et al. Does intraspecific variation in rDNA copy number affect analysis of microbial communities? *Trends Microbiol* 2021;**29**:19–27.
- Leibold MA, Holyoak M, Mouquet N et al. The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* 2004;**7**:601–13.
- Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* 2011;**9**:119–30.
- Lennon JT, Muscarella ME, Placella SA et al. How, when, and where relic DNA affects microbial diversity. *mBio* 2018;**9**:131284.
- Liu T, Zhang AN, Wang J et al. Integrated biogeography of planktonic and sedimentary bacterial communities in the Yangtze River. *Microbiome* 2018;**6**:16.
- Locey KJ, Muscarella ME, Larsen ML et al. Dormancy dampens the microbial distance–decay relationship. *Phil Trans R Soc B* 2020;**375**:20190243.
- Logares R, Bråte J, Bertilsson S et al. Infrequent marine–freshwater transitions in the microbial world. *Trends Microbiol* 2009;**17**:414–22.
- Logue JB, Lindström ES. Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints. *ISME J* 2010;**4**:729–38.
- Louca S, Jacques SMS, Pires APF et al. High taxonomic variability despite stable functional structure across microbial communities. *Nat Ecol Evol* 2016;**1**:0015.
- Lozupone CA, Knight R. Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* 2007;**104**:11436–40.
- Ma K, Tu Q. Random sampling associated with microbial profiling leads to overestimated stochasticity inference in community assembly. *Front Microbiol* 2022;**13**:1011269.
- Mercier C, Boyer F, Bonin A et al. SUMATRA and SUMACLUSt: fast and exact comparison and clustering of sequences. 2013. <http://metabarcoding.org/sumatra> (30 June 2018, date last accessed).
- Muscarella ME, Jones SE, Lennon JT. Species sorting along a subsidy gradient alters bacterial community stability. *Ecology* 2016;**97**:2034–43.
- Nagler M, Insam H, Pietramellara G et al. Extracellular DNA in natural environments: features, relevance and applications. *Appl Microbiol Biotechnol* 2018;**102**:6343–56.
- Newton RJ, Jones SE, Eiler A et al. A guide to the natural history of freshwater lake bacteria. *Microbiol Mol Biol Rev* 2011;**75**:14–49.

- Niño-García JP, Ruiz-González C, del Giorgio PA. Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. *ISME J* 2016;**10**:1755–66.
- Oksanen J, Blanchet FG, Friendly M et al. vegan: community ecology package. R Package Version 2.5-3. 2018. <http://CRAN.R-project.org/package=vegan> (1 September 2020, date last accessed).
- Painchaud J, Therriault J-C, Legendre L. Assessment of salinity-related mortality of freshwater bacteria in the Saint Lawrence estuary. *Appl Environ Microbiol* 1995;**61**:205–8.
- Paver SF, Muratore D, Newton RJ et al. Reevaluating the salty divide: phylogenetic specificity of transitions between marine and freshwater systems. *mSystems* 2018;**3**:e00232–18.
- Pedrós-Alió C. The rare bacterial biosphere. *Annu Rev Mar Sci* 2012;**4**:449–66.
- Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;**26**:1641–50.
- Quast C, Pruesse E, Yilmaz P et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:590–6.
- Quiroga MV, Valverde A, Mataloni G et al. The ecological assembly of bacterial communities in Antarctic wetlands varies across levels of phylogenetic resolution. *Environ Microbiol* 2022;**24**:3486–99.
- R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2018. <http://www.R-project.org/>.
- Raymond PA, Bauer JE. Riverine export of aged terrestrial organic matter to the North Atlantic Ocean. *Nature* 2001;**409**:497–500.
- Read DS, Gweon HS, Bowes MJ et al. Catchment-scale biogeography of riverine bacterioplankton. *ISME J* 2015;**9**:516–26.
- Richa K, Balestra C, Piredda R et al. Distribution, community composition, and potential metabolic activity of bacterioplankton in an urbanized Mediterranean Sea coastal zone. *Appl Environ Microbiol* 2017;**83**:e00494–17.
- Savio D, Sinclair L, Ijaz UZ et al. Bacterial diversity along a 2600 km river continuum. *Environ Microbiol* 2015;**17**:4994–5007.
- Schloss PD. Rarefaction is currently the best approach to control for uneven sequencing effort in amplicon sequence analyses. *mSphere* 2024;**9**:e00354–23.
- Shen D, Langenheder S, Jürgens K. Dispersal modifies the diversity and composition of active bacterial communities in response to a salinity disturbance. *Front Microbiol* 2018;**9**:2188.
- Sia ESA, Zhu Z, Zhang J et al. Biogeographical distribution of microbial communities along the Rajang River–South China Sea continuum. *Biogeosciences* 2019;**16**:4243–60.
- Smith EM, del Giorgio PA. Low fractions of active bacteria in natural aquatic communities? *Aquat Microb Ecol* 2003;**31**:203–8.
- Sogin ML, Morrison HG, Huber JA et al. Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc Natl Acad Sci USA* 2006;**103**:12115–20.
- Souffreau C, Pecceu B, Denis C et al. An experimental analysis of species sorting and mass effects in freshwater bacterioplankton. *Freshw Biol* 2014;**59**:2081–95.
- Stadler M, del Giorgio PA. Terrestrial connectivity, upstream aquatic history and seasonality shape bacterial community assembly within a large boreal aquatic network. *ISME J* 2022;**16**:937–47.
- Stegen JC, Fredrickson JK, Wilkins MJ et al. Groundwater–surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. *Nat Commun* 2016;**7**:11237.
- Stegen JC, Lin X, Fredrickson JK et al. Estimating and mapping ecological processes influencing microbial community assembly. *Front Microbiol* 2015;**6**:370.
- Stegen JC, Lin X, Fredrickson JK et al. Quantifying community assembly processes and identifying features that impose them. *ISME J* 2013;**7**:2069–79.
- Stegen JC, Lin X, Konopka AE et al. Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* 2012;**6**:1653–64.
- Stoddard SF, Smith BJ, Hein R et al. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res* 2015;**43**:D593–8.
- Sun J, Lin B, Li K et al. A modelling study of residence time and exposure time in the Pearl River Estuary, China. *J Hydro-environ Res* 2014;**8**:281–91.
- Telesh IV, Khlebovich VV. Principal processes within the estuarine salinity gradient: a review. *Mar Pollut Bull* 2010;**61**:149–55.
- Thompson LR, Sanders JG, McDonald D et al. A communal catalogue reveals Earth’s multiscale microbial diversity. *Nature* 2017;**551**:457–63.
- Urvoy M, Gourmelon M, Serghine J et al. Free-living and particle-associated bacterial community composition, assembly processes and determinants across spatiotemporal scales in a macrotidal temperate estuary. *Sci Rep* 2022;**12**:13897.
- Vellend M. Conceptual synthesis in community ecology. *Q Rev Biol* 2010;**85**:183–206.
- Wang Y, Pan J, Yang J et al. Patterns and processes of free-living and particle-associated bacterioplankton and archaeoplankton communities in a subtropical river-bay system in South China. *Limnol Oceanogr* 2020;**65**:S161–79.
- Wisnoski NI, Muscarella ME, Larsen ML et al. Metabolic insight into bacterial community assembly across ecosystem boundaries. *Ecology* 2020;**101**:e02968.
- Wu W, Liu H. Disentangling protist communities identified from DNA and RNA surveys in the Pearl River–South China Sea Continuum during the wet and dry seasons. *Mol Ecol* 2018;**27**:4627–40.
- Wu W, Xu Z, Dai M et al. Homogeneous selection shapes free-living and particle-associated bacterial communities in subtropical coastal waters. *Divers Distrib* 2021;**27**:1904–17.
- Xenopoulos MA, Downing JA, Kumar MD et al. Headwaters to oceans: ecological and biogeochemical contrasts across the aquatic continuum. *Limnol Oceanogr* 2017;**62**:S3–14.
- Zhang H, Wang K, Shen L et al. Microbial community dynamics and assembly follow trajectories of an early-spring diatom bloom in a semienclosed bay. *Appl Environ Microbiol* 2018;**84**:e01000–18.
- Zhou J, Ning D. Stochastic community assembly: does it matter in microbial ecology? *Microbiol Mol Biol Rev* 2017;**81**:e00002–17.
- Zutic V, Legovic T. A film of organic matter at the fresh-water/sea-water interface of an estuary. *Nature* 1987;**328**:612–4.