## ORIGINAL ARTICLE

# Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities

Wenxue Wu[1], Hsiao-Pei Lu[1], Akash Sastri[2], Yi-Chun Yeh[3], Gwo-Ching Gong[4], Wen-Chen Chou[4] and Chih-Hao Hsieh[1,5,6,7]

[1]Institute of Oceanography, National Taiwan University, Taipei, Taiwan; [2]Biology Department and Ocean Networks Canada, University of Victoria, Victoria, British Columbia, Canada; [3]Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA; [4]Institute of Marine Environment and Ecology, National Taiwan Ocean University, Keelung, Taiwan; [5]Department of Life Science, Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan; [6]Research Center for Environmental Changes, Academia Sinica, Taipei, Taiwan and [7]National Center for Theoretical Sciences, Taipei, Taiwan

A central challenge in microbial ecology is to understand the underlying mechanisms driving community assembly, particularly in the continuum of species sorting and dispersal limitation. However, little is known about the relative importance of species sorting and dispersal limitation in shaping marine microbial communities; especially, how they are related to organism types/traits and water depth. Here, we used variation partitioning and null model analysis to compare mechanisms driving bacterial and protist metacommunity dynamics at the basin scale in the East China Sea, based on MiSeq paired-end sequencing of 16S ribosomal DNA (rDNA) and 18S rDNA, respectively, in surface, deep chlorophyll maximum and bottom layers. Our analyses indicated that protist communities were governed more strongly by species sorting relative to dispersal limitation than were bacterial communities; this pattern was consistent across the three-depth layers, albeit to different degrees. Furthermore, we detected that bacteria exhibited wider habitat niche breadths than protists, whereas, passive dispersal abilities were not appreciably different between them. Our findings support the 'size-plasticity' hypothesis: smaller organisms (bacteria) are less environment filtered than larger organisms (protists), as smaller organisms are more likely to be plastic in metabolic abilities and have greater environmental tolerance.
*The ISME Journal* (2018) **12,** 485–494; doi:10.1038/ismej.2017.183; published online 10 November 2017

## Introduction

Microbial biogeography, which has greatly advanced because of the rapid progress in DNA sequencing technologies, concerns the distribution of microorganisms across space, and aims to identify underlying mechanisms, which generate variation in community composition (that is, beta diversity) (Martiny *et al.*, 2006). As one of the most common biogeographic patterns, the distance-decay of community similarity (Nekola and White, 1999) suggests potential driving forces underlying community assembly (Nemergut *et al.*, 2013). From the metacommunity perspective (Leibold *et al.*, 2004; Winegardner *et al.*, 2012), the strength of distance-decay of community similarity is influenced by both selective and non-selective processes (Vellend, 2010). More specifically, species sorting, dispersal limitation (alongside drift), mass effect and drift (acting alone) impose effects jointly on the distance-decay pattern (Stegen *et al.*, 2013; Bahram *et al.*, 2016). Among these, mass effect likely has a minor role in the assembly of microbial communities separated by relatively long distances (for example, marine biogeography at basin scale in this study). Thus, species sorting and dispersal limitation are the two major processes that determine the distance-decay pattern when communities are surveyed at a large spatial scale with long between-site distances (Hanson *et al.*, 2012), whereas drift weakens the distance-decay pattern via homogenizing communities.

In a relative sense, the importance of species sorting versus dispersal limitation may vary, depending on types or traits of organisms (Cottenie, 2005; Astorga *et al.*, 2012; Ragon *et al.*, 2012; Padial

486

et al., 2014). For instance, De Bie et al. (2012) showed that body size and dispersal mode were two key characteristics determining metacommunity structures of pond organisms (ranging from bacteria to fish). Soininen et al. (2011) showed that prokaryotic (bacteria) and eukaryotic (phytoplankton and zooplankton) plankton communities in 100 small Finnish lakes were differently assembled by environmental- and dispersal-driven processes. Linking organism types and/or traits to the relative importance of species sorting and dispersal limitation is therefore essential for a deeper understanding of community assembly, and may provide insights into how environmental factors act on biotic communities in the context of climate change (Soininen, 2014). In this study, we focus on marine bacterial and protist communities, which represent two pivotal components in the marine microbial food web (DeLong and Karl, 2005; Pomeroy et al., 2007). Although previous studies have explored metacommunity dynamics for marine microorganisms (Langenheder and Ragnarsson, 2007; Wang et al., 2015; Yeh et al., 2015; Wu et al., 2017), the extent to which species sorting and dispersal limitation drive patterns at different trophic levels or organism types has not been adequately explored. Importantly, as bacteria and protists have relatively comparable attributes (for example, small body size and short generation time), examining community assemblies of bacteria and protists should yield a better resolved linking of metacommunity dynamics to differences in organism attributes.

Despite their comparative traits (Massana and Logares, 2013), bacteria and protists, however, exhibit fine differences in body size, metabolic activity and dispersal potential (DeLong et al., 2010; Keeling and del Campo, 2017), which may exert an influence on the relative importance of species sorting and dispersal limitation. Two particular hypotheses, the 'size-plasticity' and 'size-dispersal' hypotheses, emphasize metabolic strategies and dispersal abilities, respectively, in structuring communities composed of different sizes of organisms (Farjalla et al., 2012). The size-plasticity hypothesis argues that smaller organisms are less environment filtered than larger organisms, because smaller organisms are more likely to have plasticity in metabolic abilities, and therefore exist widely in diverse habitats (Finlay, 2002; Langenheder et al., 2005). In contrast, the size-dispersal hypothesis predicts that smaller organisms are more likely affected by species sorting than dispersal limitation, because smaller organisms can disperse almost everywhere (Cottenie, 2005; Beisner et al., 2006; Shurin et al., 2009) and thus only reflect the environmental effects (Farjalla et al., 2012). These two hypotheses were initially explored for contrasting microscopic and macroscopic organisms; however, we suggest that the underlying ecological principles can be applied to bacterial and protist community assemblies, specifically when considering differences in metabolic strategies and body sizes of bacteria versus protists. For example, bacteria have a vast metabolic versatility, whereas protists are metabolically less flexible (Massana and Logares, 2013), and bacteria are generally smaller than protists (Kirchman, 2016). Therefore, according to the size-plasticity hypothesis, the strength of species sorting in bacterial communities, compared with that of protist communities, may be weaker because of their stronger metabolic plasticity. In contrast, and in view of the size-dispersal hypothesis, the strength of species sorting in bacterial communities can be stronger because bacteria (of smaller size) have higher dispersal abilities than protists, which weakens the role of dispersal limitation.

In this study, our aim is to compare the relative importance of species sorting and dispersal limitation in structuring bacterial versus protist communities. We carried out a three-dimensional survey by sampling the surface (SUR), deep chlorophyll maximum (DCM) and bottom (BOT) layers in the East China Sea (ECS). Importantly, previous studies concerning marine metacommunity dynamics have largely ignored the three-dimensional nature, although metacommunities are structured by either vertical or horizontal directions in the ocean (Zinger et al., 2011; de Vargas et al., 2015; Sunagawa et al., 2015). We specifically designed a sampling scheme to collect bacteria and protists simultaneously from the same set of locations, across a large environmental gradient and water masses. Simultaneous sampling is crucial for comparative studies, because the relative strength of species sorting and dispersal limitation varies across environmental conditions (Yeh et al., 2015). Bacterial and protist communities were determined with MiSeq paired-end sequencing for 16S ribosomal DNA (rDNA) and 18S rDNA, respectively. These community data were then applied to test the size-plasticity and size-dispersal hypotheses for metacommunity dynamics of bacteria versus protists. We examined the idea that differences in organism traits, such as metabolic strategies and dispersal abilities, matter for the relative strength of species sorting and dispersal limitation between microorganism groups.

## Materials and methods

*Sample collection and environmental variables*
Our sampling nearly covered the entire ECS basin (Supplementary Figure S1). Pairwise distances between sampling sites ranged from 51 to 826 km. Our sampling design with a between-site distance over 50 km is likely sufficient for capturing a clear microbial biogeography in the open ocean (Hewson et al., 2006), given the large size of a typical coherent water parcel (within which communities are relatively consistent) in marine ecosystems (Fuhrman et al., 2015). Moreover, our sampling design acknowledges

the issue that biogeographic patterns and the interaction between species sorting and dispersal limitation are dictated by the spatial scale of the observations (Rahbek, 2005; Soininen et al., 2011).

A total of 74 water samples were collected during 20–31 August 2014, using Go-Flo bottles on a CTD-equipped rosette (Sea-Bird Electronics, Bellevue, WA, USA). Sampling targeted the following three-depth layers, which encompassed distinct physico-chemical conditions: SUR (33 samples), DCM (33 samples) and BOT (19 samples). BOT samples were only collected from stations with BOT depths < 104 m (Supplementary Figure S1 and Supplementary Table S1); we did not include BOT samples from stations with greater BOT depths (for example, Station K with a depth of 1686 m; Supplementary Figure S1) into the BOT metacommunity, because those communities are not well connected with other shallower BOT samples. We also note that the SUR and DCM overlapped in 11 samples, as the chlorophyll maximum coincided with the SUR water layer at these stations. All SUR samples were taken from 5-m depth, whereas the DCM and BOT samples were collected at varying depths ranging from 5 to 80 m and 15 to 100 m, respectively, depending on the oceanography and BOT depth of the station. For each sampling, a 10–20 l water sample was continuously filtered through two size fractions (1.2 and 0.2 µm) using 142-µm polycarbonate membranes (Millipore, Billerica, MA, USA) for a coarse separation of protist and bacterial cells. Filtering was completed within 4 h of each sampling. Filters were frozen in liquid nitrogen onboard and stored at − 20 °C until DNA extraction.

Temperature and salinity were measured using a CTD profiler. Chlorophyll a, nutrients ($NH_4$, $NO_2$, $NO_3$, $PO_4$ and $SiO_3$), and dissolved oxygen were measured at each sampling depth according to standard methods (Gong et al., 2000), whereas dissolved inorganic carbon, total alkalinity, and field pH measurements were carried out as described by Chou et al. (2016).

### DNA sequencing and sequence processing

The V5–V6 region of bacterial 16S rDNA (Cai et al., 2013) and V4 region of protist 18S rDNA (Stoeck et al., 2010) were separately amplified using the total environmental DNA (see Supplementary Methods). Sequencing was performed using the v3 chemistry on a MiSeq producing 2 × 300 bp paired-end read configuration (Illumina, San Diego, CA, USA). The sequence data were processed using the Quantitative Insights Into Microbial Ecology (QIIME v. 1.9.1) pipeline (Caporaso et al., 2010) (see Supplementary Methods). For both 16S rDNA and 18S rDNA data, quality-filtered reads were clustered into operational taxonomic units (OTUs) at 97% identity using Sumaclust (Mercier et al., 2013). Further analyses based on the 95% and 99% cut-offs were performed in parallel to examine the robustness of our findings at different taxonomic resolutions (see Supplementary Results) (Martin et al., 2016). The representative sequences of OTUs were assigned at a phylum-level taxonomic rank using BLAST (E-value = $10^{-6}$) against the SILVA 119 (Quast et al., 2013) (see Supplementary Results and Supplementary Figure S2a) and $PR^2$ database (Guillou et al., 2013) (see Supplementary Results and Supplementary Figure S2b). The raw sequence data have been deposited in the NCBI Sequence Read Archive under the accession numbers PRJNA378895 and PRJNA378896.

### Community variation over environmental and spatial gradients

To assess the distance-decay of community similarity, we compared Bray–Curtis similarity (1—Bray–Curtis dissimilarity) and geographic distance matrices. The Bray–Curtis similarities were based on bacterial and protist OTU tables subsampled 100 times (hereafter, bootstrap, sensu Yeh et al., 2015) to a minimum number of sequences from each sample (67 117 for bacterial communities and 41 966 for protist communities; Supplementary Table S1). Geographic distances were determined based on the latitude and longitude of each station. We related the geographic distances and the Bray–Curtis similarities by linear regression. Significance of the relationship was tested using distance-based redundancy analysis (Legendre and Anderson, 1999), which is more reliable at analyzing community variation than Mantel's test (McArdle and Anderson, 2001).

### Variation partitioning and null model analysis

To tease apart the relative importance of species sorting and dispersal limitation, we performed variation partitioning (two-way permutational multivariate analyses of variance; McArdle and Anderson, 2001) and null model analysis (Stegen et al., 2013). Variation partitioning focuses on decomposing beta diversity of the whole metacommunity in the region, whereas the null model analysis examines the pairwise metacommunities. These two frameworks provide complementary information (Logue et al., 2011; Meynard et al., 2013; Vellend et al., 2014; Langenheder et al., 2017), and it is difficult to evaluate which one is better.

Variation partitioning was carried out to partition the community variation (pairwise Bray–Curtis dissimilarity) into environmental effects (principal component analysis of environmental variables) and spatial effects (using Moran's eigenvector maps) (Wu et al., 2017); for the environmental and spatial effects, the numbers of axes were determined according to the Kaiser–Guttman rule (see scripts in Supplementary File S1). The pure environmental variation without a spatial component ([E | S]) represents the strength of species sorting; the pure spatial variation without an environmental component

([S|E]) is interpreted as the effect of dispersal limitation. The explained variance fractions are based on adjusted fractions, which account for the number of variables and sample sizes. The significance of each component via partitioning was evaluated with the permutation test, except for the interaction component ([E∩S]) and residuals, which cannot be tested statistically (Legendre, 2008). To evaluate the relative importance of species sorting versus dispersal limitation, we compared the ratio of [E|S]/[S|E] (that is, sorting/dispersal effect ratio), rather than their absolute values. The comparison, based on the ratio, can distinguish the dominance of species sorting and dispersal limitation in a relative sense. However, comparing absolute values of these two components between different types of organisms is problematic, because cross-system comparison assumes that the amount of noise (Stegen *et al.*, 2013) and sizes of species pool (Kraft *et al.*, 2011) are similar between systems, both of which cannot hold in our data sets.

Null model analysis was carried out using the framework of Stegen *et al.* (2013) (see scripts in Supplementary File S2) to classify community pairs into underlying driving forces of species sorting, dispersal limitation, mass effect and drift. The analyses were performed using a single rarefied OTU table for both bacterial (a minimum of 67 117 sequences) and protist (41 966 sequences) communities, and the null model expectation was generated using 999 randomizations. To make variation partitioning and null model analysis comparable, we also estimated the sorting/dispersal effect ratio using the percent of pairwise communities governed by species sorting/the percent of pairwise communities governed by dispersal limitation. All of the statistical analyses described above were implemented with the R program (R Core Team, 2014).

*Habitat niche breadth*
Niche breadth is a crucial trait that influences the relative importance of species sorting and dispersal

limitation affecting communities (Pandit *et al.*, 2009). An organism group with wider niche breadth can be expected to be more metabolically flexible at the community level. Niche breadth was calculated according to Pandit *et al.* (2009), using Levins' niche breadth index (*B*):

$$B_j = 1 / \sum_{i=1}^{N} P_{ij}^2$$

where $B_j$ represents the habitat niche breadth of OTU *j* in a metacommunity; *N* is the total number of communities of each metacommunity; $P_{ij}$ is the proportion of OTU *j* in community *i*. A high *B* indicates that the OTU occurs widely and evenly along a wide range of locations, representing wide habitat niche breadth. We calculated the average *B*-values from all taxa in a single community (*B*com) as an indicator of habitat niche breadth at the community level. To contrast the overall difference in the bulk *B*com values for bacterial and protist communities, we used a generalized linear mixed-effect model (GLMM) using depth layer as a random effect with the lme4 package for R (Bates *et al.*, 2015). In addition, differences of *B*com values between bacterial and protist communities in each layer were provided and tested with analysis of variance.

*Dispersal ability*
Dispersal ability is also a critical trait that influences differences in the relative effects of species sorting and dispersal limitation (De Bie *et al.*, 2012). Although bacteria and protists are likely to be ubiquitous, we suspect that bacteria may have a greater dispersal capacity than protists because of their smaller body sizes, as can be partly supported by observations that protists are less likely to be globally dispersed (Bass *et al.*, 2007). To estimate dispersal ability of each taxon (passively driven by water movements), we calculated the pairwise shared proportion of sequence numbers, and used
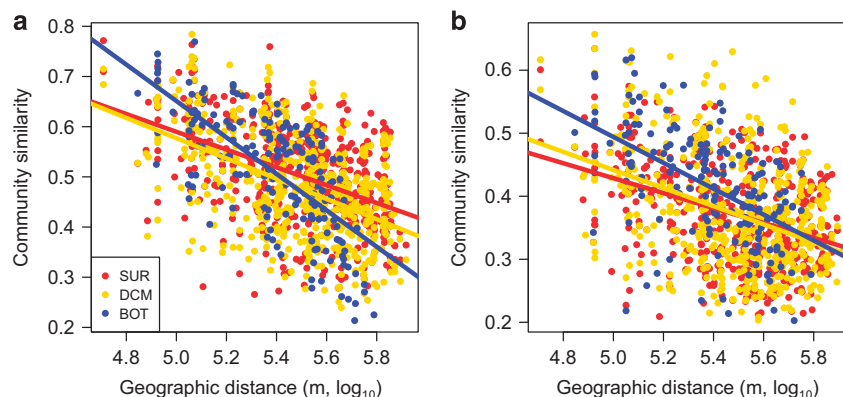


**Figure 1** Distance-decay curves showing Bray–Curtis similarity (based on the average of 100 bootstraps) between pairs of communities (**a**, bacteria; **b**, protists) against geographic distances (log$_{10}$ transformed) between sampling sites. Solid lines indicate significant correlation tested by distance-based redundancy analysis. The results are based on a 97% sequence similarity cut-off.

the average shared proportion as a proxy for dispersal ($D$). A higher shared proportion of sequences indicates a more successful passive transport by currents (Yeh et al., 2015). Again, the overall difference in the bulk $D$-values for bacteria and protists were tested by GLMM, using depth layer as a random effect. In addition, results in each layer were provided with additional comparisons (analysis of variance).

## Results

### Distance-decay of community similarity
We observed a consistent decay of community similarity with geographic distance for both bacterial and protist communities in SUR, DCM and BOT (Figure 1). The slopes of distance-decay curves were steeper for bacteria (0.18, SUR; 0.2, DCM; 0.36, BOT) than those for protists (0.12, SUR; 0.14, DCM; 0.21, BOT) in all three-depth layers. Environmental variables were characterized by strong inshore–offshore gradients (Supplementary Figure S3), which were significantly related to community dissimilarities (Supplementary Figure S4).

### Quantifying the relative effects of species sorting and dispersal limitation
The sorting/dispersal effect ratios (Figure 2) were consistently higher in protist communities regardless of the depth layer (variation partitioning: 0.43, SUR; 0.52, DCM; 1.91, BOT; null model: 46, SUR; 5.47, DCM; 76) than in bacterial communities (variation partitioning: 0.17, SUR; 0.22, DCM; 0.5, BOT; null model: 1.39, SUR; 0.81, DCM; 2.25). According to variation partitioning, all the pure effects of environmental and spatial variables (that is, [E|S] and [S|E]) were significant ($P<0.05$) (Supplementary Figure S5), except for the pure environmental fraction in the DCM for bacteria (Supplementary Figure S5b). Null model analysis revealed a negligible proportion (0%) of dispersal effect for the BOT protists (Supplementary Figure S6). These results

indicated that protist communities were more governed by species sorting relative to dispersal limitation than bacterial communities.

To compare metacommunity dynamics across SUR, DCM and BOT, we further carried out analyses using only the common stations (19 stations where samples were taken in all three layers; Supplementary Figure S1). These analyses for the subset of common stations showed patterns similar to the findings based on the whole data set, indicating that our conclusion is not biased by the unequal sample sizes among depth layers (Supplementary Figure S7). For variation partitioning, we found that for bacteria, the effect of dispersal limitation increased with increasing depth (Supplementary Figures S7a–c); by contrast, for protists, the effect of species sorting increased with increasing depth (Supplementary Figures S7d–f). Overall, the differences in sorting/dispersal effect ratio between bacteria and protists became larger with increasing depth (bacteria/protists: 0.71/1.24 at SUR, 0.3/1.36 at DCM, 0.5/1.91 at BOT; Supplementary Figure S9). Note, similar comparison of sorting/dispersal effect ratios based on the null model analysis could not be achieved, because we found that 0% of protist community pairs were assembled by dispersal limitation in both SUR and BOT (Supplementary Figure S8).

### Niche breadth and dispersal ability
The bulk community habitat niche breadths ($B$com) for bacterial communities were significantly higher than that of protist communities (GLMM, inclusive of all three layers, $P<0.001$) (Figure 3). In all three layers, the mean $B$com values for bacterial communities (6.61, SUR; 6.1, DCM; 4.81, BOT) were higher than that for protist communities (4.71, SUR; 4.79, DCM; 4.48, BOT).

The bulk dispersal abilities ($D$) for bacterial communities were slightly higher than that of protist communities (GLMM, inclusive of all three layers, $P<0.001$) (Figure 4). Bacterial communities (46.7%,
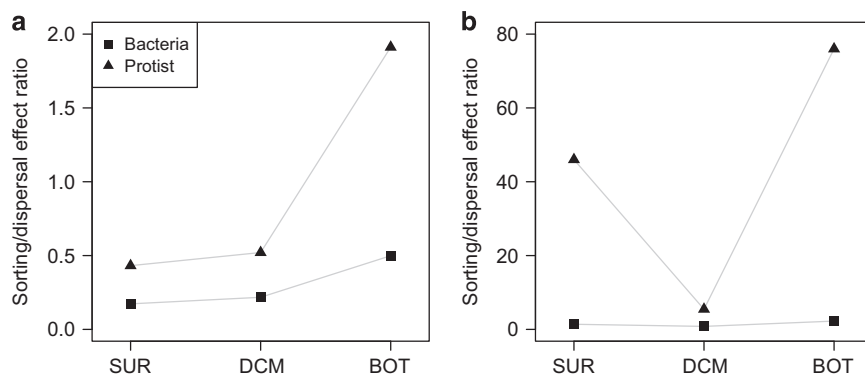
**Figure 2** The sorting/dispersal effect ratio (**a**, variation partitioning; **b**, null model analysis) of bacterial versus protist communities in SUR, DCM and BOT layers. For null model analysis, in the case (protist communities in BOT) where the proportion of community pairs assembled by dispersal limitation was 0%, we recalculated the sorting/dispersal effect ratio using a proportion of 0.6%, which represents only 1 pair of communities classified as dispersal limitation. The results are based on a 97% sequence similarity cut-off.
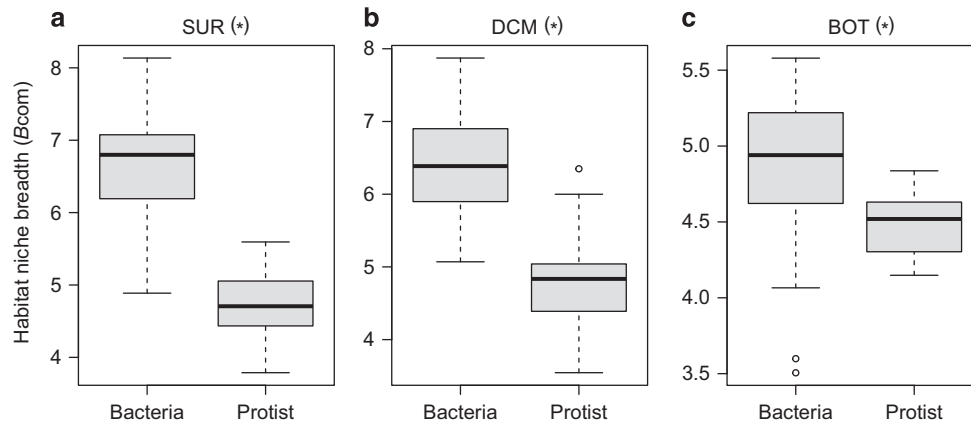
490



**Figure 3** Boxplots illustrating mean habitat niche breadth from all taxa (based on a 97% sequence similarity) in each sample (*B*com) of bacterial versus protist communities in (**a**) SUR, (**b**) DCM and (**c**) BOT layers. The bulk *B*com values for bacteria are significantly higher than that of protists (GLMM with layer as random effect, $P < 0.001$). The significant difference between bacterial versus protist communities in each layer is further marked with an asterisk in brackets ($P < 0.05$).
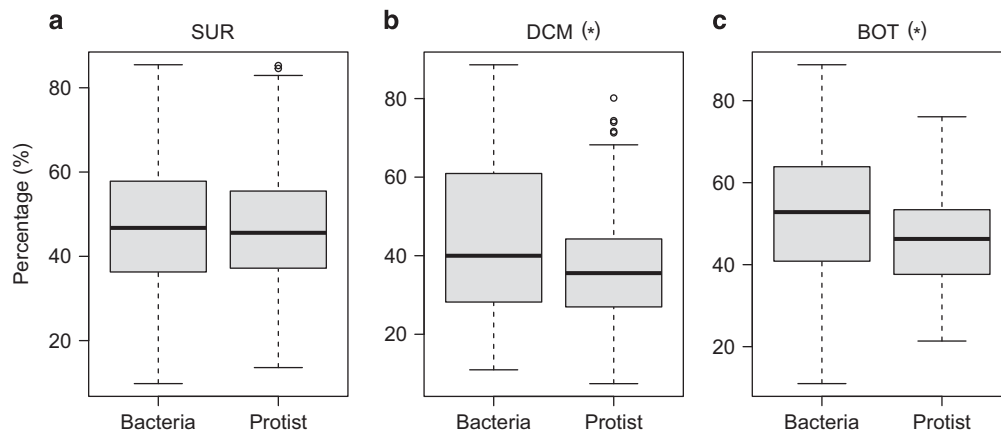


**Figure 4** Boxplots showing mean shared proportions (%) of sequences of bacterial and protist communities (based on a 97% sequence similarity) in (**a**) SUR, (**b**) DCM and (**c**) BOT layers. A greater proportion indicates greater dispersal ability of organisms in general. The bulk shared proportions for bacteria are significantly higher than that of protists (GLMM with layer as random effect, $P < 0.001$). The significant difference between bacterial versus protist communities in each layer is further marked with an asterisk in brackets ($P < 0.05$).

SUR; 44.2%, DCM; 50.8%, BOT) showed significantly higher $D$ than protist communities (46%, SUR; 36.3%, DCM; 45.8%, BOT) at the DCM and BOT but not at the SUR; however, the differences were not substantial.

Qualitatively similar results were obtained when considering the 19 common stations (Supplementary Figure S1) for the three-depth layers (GLMM with layer as random effect, $P < 0.001$). That is, bacteria generally had much wider habitat niche breadths than protists (Supplementary Figure S10), and bacteria exhibited greater dispersal abilities than protists, although the difference is not that substantial (Supplementary Figure S11).

As a sensitivity test, we repeated the analyses based on 95% and 99% sequence similarity cut-offs. The results of sorting/dispersal effect ratio, habitat niche breadth and dispersal ability at 95% and 99% sequence similarity cut-offs were consistent with those based on the 97% cut-off (Supplementary Figure S12).

## Discussion

*Community assembly of bacteria versus protists*
In this study, we detected lower sorting/dispersal effect ratios for bacterial communities relative to protist communities in all three-depth layers (Figure 2), indicating that protist communities were more governed by species sorting relative to dispersal limitation than were bacterial communities. This pattern supports the size-plasticity hypothesis: smaller organisms (bacteria) are less environment filtered than larger organisms (protists) (Farjalla *et al.*, 2012).

To explain this pattern, we quantified the community-level habitat niche breadth (Figure 3) and dispersal ability (Figure 4) for bacteria versus protists in the ECS. We found that bacteria generally had wider niche breadths than protists; this finding was in agreement with Pandit *et al.* (2009), indicating that habitat generalists with wider niche breadths were less influenced by environmental factors. In

fact, the wider niche breadth of bacteria might imply greater metabolic plasticity; for example, some marine bacteria can even survive in freshwater environments (Comte et al., 2014).

An alternative explanation for this pattern may be related to dormancy strategies, which are more commonly observed in bacteria than protists (Jones and Lennon, 2010; Massana and Logares, 2013). Dormancy is an essential metabolic strategy for microbial cells to cope with environmental stresses, as dormant cells are likely to be less susceptible to selection processes (Lennon and Jones, 2011; Nemergut et al., 2013). For bacteria in various environments, $<10\%$ of the cells in a given community are in an active stage at any time (Locey, 2010). In fact, marine bacterioplankton are often found to be dormant (Gasol et al., 1995). By contrast, a dormancy strategy might have a minor role for protists (Jones and Lennon, 2010), although protists do have the ability to form dormant stages (for example, cysts, eggs and spores). It is possible that protist communities are more responsive to species sorting than bacterial communities, because protists have a more limited tendency to enter dormancy (Massana and Logares, 2013). However, it is difficult to distinguish the effects of dormancy and the aforementioned metabolic plasticity on community assembly, as the wider niche breadth of bacterial taxa may also result from their greater potential for dormancy.

In addition to wider niche breadth, we found that bacteria have slightly higher dispersal abilities than protists (Figure 4); however, this difference is not substantial, compared with the large difference in niche breadth (Figure 3). More importantly, our results do not support the size-dispersal hypothesis; that is, bacteria, even with slightly higher dispersal abilities, do not reflect more environmental influence than protists in the ECS.

A few caveats should be considered in the interpretation of our findings which contrast the community assembly of bacteria versus protists. First, one may argue that a large proportion of variation remains unexplained for both bacterial and protist communities in the variation partitioning (Supplementary Figures S5 and S12). The unexplained variation may be largely attributed to unmeasured environmental variables. However, the unexplained variation in our study is in fact lower than that in previous studies (for example, Langenheder and Ragnarsson, 2007; Hájek et al., 2011; De Bie et al., 2012; Farjalla et al., 2012). Second, the drift that results from stochastic differences in population size, birth and death rates (Stegen et al., 2013) may exhibit differential influences on bacterial and protist metacommunity dynamics (Kirchman, 2016). For instance, bacteria generally have greater population sizes than protists and are therefore, generally less influenced by drift (at 95% and 97% cut-offs; Supplementary Figures S6 and S13), which may then influence the strength of

dispersal limitation that operates alongside drift (Stegen et al., 2013). In addition, we acknowledge that biotic interactions are omitted in our analyses, although these forces (for example, prey–predator interaction, competition) are important as a part of species sorting (Horner-Devine et al., 2007; Jeganmohan et al., 2014; Berga et al., 2015). For instance, differences in selective grazing and trophic levels between prey (bacteria) and predator (protists) may obscure our ability to reliably quantify species sorting effects (Livingston et al., 2017).

As a note, we wish to emphasize the rationale of using the sorting/dispersal effect ratio, rather than the absolute values accounting for sorting and dispersal, to evaluate the relative importance of species sorting and dispersal limitation when comparing different types of organisms. If we had only compared the absolute values of dispersal effect between protist and bacterial communities (Supplementary Figure S5), we would conclude that the bacterial communities were more dispersal limited than the protist communities; this conclusion contradicts the finding based on dispersal capacity (Figure 4). Logically, it is impossible to conclude that bacteria simultaneously have higher dispersal capacities and experience stronger dispersal limitation. Thus, we made the comparison based on the sorting/dispersal effect ratio, instead of absolute values representing species sorting and dispersal limitation in variation partitioning and null model analysis.

*Variation of metacommunity dynamics across depths*
Based on the analyses of the 19 common stations, we found that, with increasing depth (SUR versus BOT), protists were increasingly affected by species sorting relative to dispersal limitation; that is, the difference in the sorting/dispersal effect ratio between bacteria and protists became larger with depth (Supplementary Figures S9 and S19). Two scenarios may explain this contrast: first, the effect of dispersal limitation for bacteria increased with depth (relative to protists), and/or, second, the effect of species sorting for protists increased with depth (relative to bacteria). These two scenarios need not be mutually exclusive. Notably, owing to the overlap of the SUR and DCM samples (8 out of 19; Supplementary Figure S1), we only consider SUR and BOT in comparison across depths.

In the scenario 1, the pure effect of dispersal limitation in bacterial communities increased considerably from SUR to BOT (for example, Supplementary Figures S7a, S7c, S8a and S8c). Considering patch connectivity driven by wind dispersal (Horváth et al., 2016), this pattern indicated that the degree of patch connectivity among ECS sampling sites decreased with depth, because of the effect of wind, decaying from the shallow to deep layers. In contrast, the pure effect of dispersal limitation was not vertically enhanced with depth

in protist communities (for example, Supplementary Figures S7d, S7f, S8a and S8c). We speculate that patch connectivity regarding dispersal limitation may have a less critical role in the vertical variation of protist communities, because of their weaker dispersal capability (Supplementary Figure S11).

In scenario 2, the pure effect of species sorting in protist communities increased from SUR to BOT (for example, Supplementary Figures S7d, S7f, S8a and S8c), which is mainly supplied by factors including BOT depth, $O_2$ and $NO_2$ (Supplementary Figure S4f; Supplementary Table S2). We suggest that the BOT depth is likely a proxy representing how strongly the hydrographic conditions in the BOT layer were influenced by sediment characteristics (Koop et al., 1990), given that all BOT communities are close to sediments (<7 m; Supplementary Table S1). The effect of $O_2$ may be related to coastal hypoxia, which has been commonly observed in the ECS during summer (Chen et al., 2007). In line with our results, Orsi et al. (2012) observed that protist communities were strongly shaped by $O_2$ gradients in an anoxic fjord in the coast of Vancouver Island British Columbia. Furthermore, $NO_2$ can act as an alternative electron acceptor in $O_2$-deficient waters, and thus the $NO_2$ patterns may influence protist communities by involving microbial nitrogen cycling processes in the $O_2$-deficient BOT (Lam and Kuypers, 2011; Orsi et al., 2012). In contrast with the pattern for protist communities, the pure effects of species sorting in bacterial communities did not increase from SUR to BOT (Supplementary Figures S7a and S7c), or exhibited a relatively modest increase (Supplementary Figures S8a and S8c); this may be related to wider habitat niche breadth (that is, higher environmental tolerances) of bacterial taxa (Supplementary Figure S10) that would complicate environment-community relationships (Farjalla et al., 2012).

Notably, the interpretation above relies on the absolute values of sorting and dispersal effects for bacteria or protist communities in different depth layers. For this comparison to make sense, the amount of noise is assumed to be consistent for communities of different layers for a given type of organism.

*Concluding remarks*
In the ECS, protist communities were governed to a greater extent by species sorting relative to dispersal limitation than bacterial communities. This result is robust across taxonomic resolutions (from 95% to 99% sequence similarity cut-offs) and depth layers (SUR, DCM and BOT). These contrasting patterns may be attributed to differences in habitat niche breadth (bacteria > protists) and dormancy tendency (bacteria > protists) between bacteria and protists. Moreover, we detected that the difference in the sorting/dispersal effect ratio between protists and bacteria became larger in deeper layers, suggesting

that the relative effect of dispersal limitation on bacterial communities increased with increasing depth, whereas the relative effect of species sorting was enhanced for protist communities. From a biogeography perspective, our results highlight the importance of considering organism characteristics and vertical structures of hydrographic conditions for studying metacommunity dynamics in marine ecosystems.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgements

## References

Astorga A, Oksanen J, Luoto M, Soininen J, Virtanen R, Muotka T. (2012). Distance decay of similarity in freshwater communities: do macro- and microorganisms follow the same rules? *Glob Ecol Biogeogr* **21**: 365–375.

Bahram M, Kohout P, Anslan S, Harend H, Abarenkov K, Tedersoo L. (2016). Stochastic distribution of small soil eukaryotes resulting from high dispersal and drift in a local environment. *ISME J* **10**: 885–896.

Bass D, Richards TA, Matthai L, Marsh V, Cavalier-Smith T. (2007). DNA evidence for global dispersal and probable endemicity of protozoa. *BMC Evol Biol* **7**: 162.

Bates D, Mächler M, Bolker B, Walker S. (2015). Fitting linear mixed-effects models using lme4. *J Stat Softw* **67**: 1–48.

Beisner BE, Peres-Neto PR, Lindström ES, Barnett A, Longhi ML. (2006). The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology* **87**: 2985–2891.

Berga M, Östman Ö, Lindström ES, Langenheder S. (2015). Combined effects of zooplankton grazing and dispersal on the diversity and assembly mechanisms of bacterial metacommunities. *Environ Microbiol* **17**: 2275–2287.

Cai L, Ye L, Tong AHY, Lok S, Zhang T. (2013). Biased diversity metrics revealed by bacterial 16S pyrotags derived from different primer sets. *PLoS ONE* **8**: e53649.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Chen C-C, Gong G-C, Shiah F-K. (2007). Hypoxia in the East China Sea: one of the largest coastal low-oxygen areas in the world. *Mar Environ Res* **64**: 399–408.

Chou W-C, Gong G-C, Yang C-Y, Chuang K-Y. (2016). A comparison between field and laboratory pH measurements for seawater on the East China Sea shelf. *Limnol Oceanogr Methods* **14**: 315–322.

Comte J, Lindström ES, Eiler A, Langenheder S. (2014). Can marine bacteria be recruited from freshwater sources and the air? *ISME J* **8**: 2423–2430.

Cottenie K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecol Lett* **8**: 1175–1182.

De Bie T, De Meester L, Brendonck L, Martens K, Goddeeris B, Ercken D *et al.* (2012). Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecol Lett* **15**: 740–747.

de Vargas C, Audic S, Henry N, Decelle J, Mahe F, Logares R *et al.* (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**: 1261605.

DeLong EF, Karl DM. (2005). Genomic perspectives in microbial oceanography. *Nature* **437**: 336–342.

DeLong JP, Okie JG, Moses ME, Sibly RM, Brown JH. (2010). Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life. *Proc Natl Acad Sci USA* **107**: 12941–12945.

Farjalla VF, Srivastava DS, Marino NAC, Azevedo FD, Dib V, Lopes PM *et al.* (2012). Ecological determinism increases with organism size. *Ecology* **93**: 1752–1759.

Finlay BJ. (2002). Global dispersal of free-living microbial eukaryote species. *Science* **296**: 1061–1063.

Fuhrman JA, Cram JA, Needham DM. (2015). Marine microbial community dynamics and their ecological interpretation. *Nat Rev Microbiol* **13**: 133–146.

Gasol JM, del Giorgio PA, Massana R, Duarte CM. (1995). Active versus inactive bacteria: size-dependence in a coastal marine plankton community. *Mar Ecol Prog Ser* **128**: 91–97.

Gong G-C, Shiah F-K, Liu K-K, Wen Y-H, Liang M-H. (2000). Spatial and temporal variation of chlorophyll *a*, primary productivity and chemical hydrography in the southern East China Sea. *Cont Shelf Res* **20**: 411–436.

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L *et al.* (2013). The protist ribosomal reference database (PR²): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* **41**: D597–D604.

Hájek M, Roleček J, Cottenie K, Kintrová K, Horsák M, Poulíčková A *et al.* (2011). Environmental and spatial controls of biotic assemblages in a discrete semi-terrestrial habitat: comparison of organisms with different dispersal abilities sampled in the same plots. *J Biogeogr* **38**: 1683–1693.

Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* **10**: 497–506.

Hewson I, Steele JA, Capone DG, Fuhrman JA. (2006). Temporal and spatial scales of variation in bacterioplankton assemblages of oligotrophic surface waters. *Mar Ecol Prog Ser* **311**: 67–77.

Horner-Devine MC, Silver JM, Leibold MA, Bohannan BJM, Colwell RK, Fuhrman JA *et al.* (2007). A comparison of taxon co-occurrence patterns for macro- and microorganisms. *Ecology* **88**: 1345–1353.

Horváth Z, Vad CF, Ptacnik R. (2016). Wind dispersal results in a gradient of dispersal limitation and environmental match among discrete aquatic habitats. *Ecography* **39**: 726–732.

Jeganmohan S, Tucker C, Cadotte MW. (2014). Colonization rates in a metacommunity altered by competition. *PLoS ONE* **9**: e88344.

Jones SE, Lennon JT. (2010). Dormancy contributes to the maintenance of microbial diversity. *Proc Natl Acad Sci USA* **107**: 5881–5886.

Keeling PJ, del Campo J. (2017). Marine protists are not just big bacteria. *Curr Biol* **27**: R541–R549.

Kirchman DL. (2016). Growth rates of microbes in the oceans. *Ann Rev Mar Sci* **8**: 285–309.

Koop K, Boynton WR, Wulff F, Carman R. (1990). Sediment-water oxygen and nutrient exchanges along a depth gradient in the Baltic Sea. *Mar Ecol Prog Ser* **63**: 65–77.

Kraft NJB, Comita LS, Chase JM, Sanders NJ, Swenson NG, Crist TO *et al.* (2011). Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science* **333**: 1755–1758.

Lam P, Kuypers MMM. (2011). Microbial nitrogen cycling processes in oxygen minimum zones. *Ann Rev Mar Sci* **3**: 317–345.

Langenheder S, Lindström ES, Tranvik LJ. (2005). Weak coupling between community composition and functioning of aquatic bacteria. *Limnol Oceanogr* **50**: 957–967.

Langenheder S, Ragnarsson H. (2007). The role of environmental and spatial factors for the composition of aquatic bacterial communities. *Ecology* **88**: 2154–2161.

Langenheder S, Wang J, Karjalainen SM, Laamanen TM, Tolonen KT, Vilmi A *et al.* (2017). Bacterial metacommunity organization in a highly-connected aquatic system. *FEMS Microbiol Ecol* **93**: fiw225.

Legendre P. (2008). Studying beta diversity: ecological variation partitioning by multiple regression and canonical analysis. *J Plant Ecol* **1**: 3–8.

Legendre P, Anderson MJ. (1999). Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecol Monogr* **69**: 1–24.

Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF *et al.* (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* **7**: 601–613.

Lennon JT, Jones SE. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* **9**: 119–130.

Livingston G, Fukumori K, Provete DB, Kawachi M, Takamura N, Leibold MA. (2017). Predators regulate prey species sorting and spatial distribution in microbial landscapes. *J Anim Ecol* **86**: 501–510.

Locey KJ. (2010). Synthesizing traditional biogeography with microbial ecology: the importance of dormancy. *J Biogeogr* **37**: 1835–1841.

Logue JB, Mouquet N, Peter H, Hillebrand H. (2011). Empirical approaches to metacommunities: a review and comparison with theory. *Trends Ecol Evol* **26**: 482–491.

Martin GK, Adamowicz SJ, Cottenie K. (2016). Taxonomic resolution based on DNA barcoding affects environmental signal in metacommunity structure. *Freshw Sci* **35**: 701–711.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al.* (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**: 102–112.

494

Massana R, Logares R. (2013). Eukaryotic versus prokaryotic marine picoplankton ecology. *Environ Microbiol* **15**: 1254–1261.

McArdle BH, Anderson MJ. (2001). Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* **82**: 290–297.

Mercier C, Boyer F, Bonin A, Coissac E. (2013). SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. Available: http://metabarcoding.org/sumatra.

Meynard CN, Lavergne S, Boulangeat I, Garraud L, Van Es J, Mouquet N *et al.* (2013). Disentangling the drivers of metacommunity structure across spatial scales. *J Biogeogr* **40**: 1560–1571.

Nekola JC, White PS. (1999). The distance decay of similarity in biogeography and ecology. *J Biogeogr* **26**: 867–878.

Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF *et al.* (2013). Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* **77**: 342–356.

Orsi W, Song YC, Hallam S, Edgcomb V. (2012). Effect of oxygen minimum zone formation on communities of marine protists. *ISME J* **6**: 1586–1601.

Padial AA, Ceschin F, Declerck SAJ, De Meester L, Bonecker CC, Lansac-Tôha FA *et al.* (2014). Dispersal ability determines the role of environmental, spatial and temporal drivers of metacommunity structure. *PLoS ONE* **9**: e111227.

Pandit SN, Kolasa J, Cottenie K. (2009). Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* **90**: 2253–2262.

Pomeroy L, Williams PJ, Azam F, Hobbie J. (2007). The microbial loop. *Oceanography* **20**: 28–33.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P *et al.* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: 590–596.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria. Available at: http://www.R-project.org/.

Ragon M, Fontaine MC, Moreira D, López-García P. (2012). Different biogeographic patterns of prokaryotes and microbial eukaryotes in epilithic biofilms. *Mol Ecol* **21**: 3852–3868.

Rahbek C. (2005). The role of spatial scale and the perception of large-scale species-richness patterns. *Ecol Lett* **8**: 224–239.

Shurin JB, Cottenie K, Hillebrand H. (2009). Spatial autocorrelation and dispersal limitation in freshwater organisms. *Oecologia* **159**: 151–159.

Soininen J. (2014). A quantitative analysis of species sorting across organisms and ecosystems. *Ecology* **95**: 3284–3292.

Soininen J, Korhonen JJ, Karhu J, Vetterli A. (2011). Disentangling the spatial patterns in community composition of prokaryotic and eukaryotic lake plankton. *Limnol Oceanogr* **56**: 508–520.

Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ *et al.* (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J* **7**: 2069–2079.

Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner H-W *et al.* (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* **19**: 21–31.

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G *et al.* (2015). Structure and function of the global ocean microbiome. *Science* **348**: 1261359.

Vellend M. (2010). Conceptual synthesis in community ecology. *Q Rev Biol* **85**: 183–206.

Vellend M, Srivastava DS, Anderson KM, Brown CD, Jankowski JE, Kleynhans EJ *et al.* (2014). Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos* **123**: 1420–1430.

Wang K, Ye X, Chen H, Zhao Q, Hu C, He J *et al.* (2015). Bacterial biogeography in the coastal waters of northern Zhejiang, East China Sea is highly controlled by spatially structured environmental gradients. *Environ Microbiol* **17**: 3898–3913.

Winegardner AK, Jones BK, Ng ISY, Siqueira T, Cottenie K. (2012). The terminology of metacommunity ecology. *Trends Ecol Evol* **27**: 253–254.

Wu W, Logares R, Huang B, Hsieh C. (2017). Abundant and rare picoeukaryotic sub-communities present contrasting patterns in the epipelagic waters of marginal seas in the northwestern Pacific Ocean. *Environ Microbiol* **19**: 287–300.

Yeh Y-C, Peres-Neto PR, Huang S-W, Lai Y-C, Tu C-Y, Shiah F-K *et al.* (2015). Determinism of bacterial metacommunity dynamics in the southern East China Sea varies depending on hydrography. *Ecography* **38**: 198–212.

Zinger L, Amaral-Zettler LA, Fuhrman JA, Horner-Devine MC, Huse SM, Welch DBM *et al.* (2011). Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS ONE* **6**: e24570.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)