Homogeneous selection shapes free-living and particle-associated bacterial communities in subtropical coastal waters

Wenxue Wu¹,²,³ | Zhimeng Xu⁴ | Minhan Dai⁵ | Jianping Gan⁴ | Hongbin Liu²,⁴

¹School of Marine Sciences, Sun Yat-sen University, Zhuhai, China
²Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China
³Guangdong Provincial Key Laboratory of Marine Resources and Coastal Engineering, Sun Yat-sen University, Zhuhai, China
⁴Department of Ocean Science, The Hong Kong University of Science and Technology, Kowloon, Hong Kong SAR, China
⁵State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen, China

Correspondence
Hongbin Liu, Department of Ocean Science, The Hong Kong University of Science and Technology, Kowloon, Hong Kong SAR, China.
Email: liuhb@ust.hk

Funding information
Theme-based Research Scheme of the Hong Kong Research Grants Council, Grant/Award Number: T21/602/16.

Abstract
Aim: In microbial biogeography, it is crucial to link spatial patterns with underlying drivers in natural ecosystems. Bacterial communities driving key biogeochemical processes in coastal zones, which are important interfaces between terrestrial and marine ecosystems, are affected by perturbations due to both natural and anthropogenic factors. However, the assembly of bacterioplankton communities, either free-living (FL) or particle-associated, in coastal ecosystems is still poorly understood.

Location: Coastal South China Sea influenced by the Pearl River (SCSPR).

Methods: In this study, we investigated FL, nanoparticle-associated (NA) and micro-particle-associated (MA) bacterial communities in the SCSPR, using environmental DNA metabarcoding based on the V4 region of the 16S rRNA gene. We assessed the relative importance of ecological processes using null model analyses based on a two-step framework.

Results: We found that the observed amplicon sequence variants (ASVs) increased from FL, NA to MA communities, and a remarkably pervasive core set of ASVs closely belonged to potential hydrocarbonoclastic bacteria with wide habitat niche breadths. Analyses of similarity tests revealed that FL, NA and MA communities differed significantly but weakly, based on unweighted ($R = 0.27, p < .001$) and weighted ($R = 0.18, p < .001$) dissimilarities. Fundamental regulation of bacterial communities via homogeneous selection was most prominent in FL (90%), followed by NA (86.8%) and MA (73.3%), whereas the relative importance of dispersal limitation was most pronounced in MA (13.2%), followed by NA (11.7%) and FL (8.6%).

Main conclusions: Bacterial communities in the SCSPR can be predominantly influenced by human activities, as indicated by the core ASVs closely related to hydrocarbonoclastic bacteria. Importantly, anthropogenic influences can temporally overwhelm the large environmental heterogeneities in coastal ecosystems. This study fills knowledge gaps in bacterial community assembly, which may facilitate future studies regarding anthropogenic influences on coastal ecosystems.

Keywords
community assembly, environmental DNA, homogeneous selection, hydrocarbonoclastic bacteria, metabarcoding
1 | INTRODUCTION

A central aspect of microbial biogeography is to link spatial patterns with underlying drivers in natural ecosystems (Hanson et al., 2012; Langenheder & Lindström, 2019; Martiny et al., 2006; Nemergut et al., 2013; Zhou & Ning, 2017). According to the metacommunity concept (Leibold et al., 2004), deterministic and stochastic processes are primarily responsible for the spatial turnover of natural communities (Vellend, 2010). The term “deterministic process” refers to two types of selective forces, namely, those that lead to either more (i.e. homogenous selection) or less (i.e. heterogeneous selection) similar structures among communities due to homogenous and heterogeneous environmental pressures, respectively (Zhou & Ning, 2017). Meanwhile, the term “stochastic process” refers to homogenizing dispersal, dispersal limitation (combined with drift) and pure drift, which can obscure the turnover among microbial communities due to high dispersal; low dispersal; and random changes in birth, death and reproduction, respectively (Zhou & Ning, 2017). The relative importance of diversification, occurring on a large spatial and temporal scale, is disregarded because it is usually minor (Stegen et al., 2013, 2015).

Marine bacterioplankton have traditionally been classified into two types of communities: free-living (FL) and particle-associated (DeLong et al., 1993), both of which play fundamental roles in marine ecosystems (Azam & Long, 2001). These two categories, however, exhibit significant differences, both in terms of composition and function (D’Ambrosio et al., 2014). For example, FL bacteria are generally smaller than their particle-associated counterparts (Dang & Lovell, 2016), whereas particle-associated bacteria exhibit higher production (Crump et al., 1998) and carbon demand (Lapoussière et al., 2011). Particle-associated bacteria are highly effective at converting high-molecular weight organic matter into smaller substrates, which are bioavailable to their FL cohorts (Arnosti, 2011; Azam & Long, 2001). The truly planktonic and sessile states of marine bacteria allow two entirely different lifestyles owing to differences in phylogenetically conserved traits (Salazar et al., 2015) in their respective genomes (Lauro et al., 2009; Smith et al., 2013). Importantly, these differences can largely mediate the balance between stochastic and deterministic processes (Xu, Zhang et al., 2020). Although many previous studies have described the community compositions of FL and particle-associated bacteria, how bacteria structure their ecological processes remains to be elucidated.

Coastal ecosystems harbour complex bacterial communities (Crump et al., 2004), which mirror the simultaneous influences of both natural and anthropogenic perturbations (Aguirre et al., 2017; Jeffries et al., 2016). One of the most salient characteristics of coastal ecosystems is the natural heterogeneity of physiochemical conditions, including light, temperature, salinity and nutrients, which are influenced by factors such as intermingling of freshwater and seawater, as well as transportation of dissolved and suspended matter. Prominent heterogeneities, such as differences in salinity, can acutely regulate bacterial communities (Lozupone & Knight, 2007). In addition, coastal ecosystems are intensely influenced by human activities, resulting in increased levels of nutrients (Cloern et al., 2016; Halpern et al., 2008) and hydrocarbons (Duran & Cravo-Laureau, 2016); these pollutants cause spatial heterogeneities in bacterial communities (Crump et al., 2004; del Giorgio & Bouvier, 2002; Kirchman et al., 2005). Therefore, it is important to improve our understanding of the bacterial communities in coastal ecosystems such as the coastal South China Sea influenced by the Pearl River (SCSPR), to reflect changes occurring in coastal waters globally (Rabalais et al., 2009). The Pearl River is the second largest river in China in terms of annual discharge (336 km$^3$), and the region around its delta is highly urbanized and industrialized (Zhang et al., 2007). As a consequence, the SCSPR is susceptible to intense perturbations caused by natural processes and anthropogenic activities (Zhang et al., 2007). This is especially obvious during the wet season, between April and September, when ~80% of the annual runoff occurs. The large input of freshwater is accompanied by heavy loading of nutrients, causing eutrophication, which is a major concern as it leads to the formation of a hypoxic bottom layer (Rabalais et al., 2009). The SCSPR is therefore a model system for studying threatened coastal ecosystems (Fennel & Testa, 2019), where changes in bacterial communities may precede detrimental effects on macroorganisms (Spitz et al., 2016).

In this study, we investigated FL, nanoparticle-associated (NA) and microparticle-associated (MA) bacterial communities at the surface and in the bottom layer of the SCSPR during the summer wet season. Bacterial communities were identified using environmental DNA metabarcoding based on the V4 region of 16S rRNA gene. We evaluated the underlying mechanisms of community assembly using null model analyses and identified a set of core bacteria, comprising potentially hydrocarbon-degrading members, which indicated acute perturbations due to anthropogenic activities. Altogether, the findings of this study provide a better understanding of the assembly of bacterial communities in subtropical coastal waters.

2 | METHODS

2.1 | Sample collection

Between 10 July and 21 July 2017, water samples for molecular analyses were collected from the surface and bottom layers at 19 stations in the SCSPR (Figure 1) using Niskin bottles mounted on an SBE 32 carousel water sampler (Sea-Bird Electronics). Seawater (350–1,000 ml) was pre-filtered through a 200 μm mesh and then sequentially through 20-, 3- and 0.2-μm-pore size polycarbonate membranes (47 mm diameter, Millipore). The size continuum of 0.2–3, 3–20 and 20–200 μm represents the collections of FL, NA and MA bacterial communities, respectively (Andersen et al., 2016; Cui et al., 2019). The filters were stored at ~20°C onboard and at ~80°C for post-cruise analyses in the laboratory. Ten MA communities were derived from filters immersed in RNA-later solution (Ambion) or amplified using additional thermal cycles (see procedures below) due to difficulties in molecular analyses (Table S1).
The temperature, salinity and turbidity were measured using a conductivity–temperature–depth (CTD) profiler (Sea-Bird Electronics). Chlorophyll \(a\) concentration was measured using 90% acetone extractions of GF/F membrane-filtered water samples with a fluorometer (Turner Designs). The concentration of dissolved oxygen was determined onboard by the Winkler titration method, using a UV-1800 spectrophotometer (Shimadzu). The numbers of cyanobacterial, heterotrophic bacterial and pigmented eukaryotic cells were counted using a FACSCalibur flow cytometer (BD Biosciences) as described by Chen et al. (2009). Other parameters, including dissolved organic carbon (Dai et al., 2009), nutrients (NO\(_3\), PO\(_4\), SiO\(_3\)) and total suspended matter, were also measured using standard protocols (Han et al., 2012). Pearson’s correlations and principal component analyses of standardized environmental variables (with a mean of 0 and a variance of 1 for each variable) were visualized using the corrplot (Wei & Simko, 2017) and vegan (Oksanen et al., 2018) packages, respectively, in R (R Core Team, 2018). Variables showing high Pearson’s correlations (\(<-0.8\) or \(>0.8\)) with other variables were excluded from the principal component analysis (i.e. temperature, dissolved organic carbon concentration, NO\(_3\), and SiO\(_3\); Figure S1).

### 2.2 DNA extraction and metabarcoding

Each filter was cut into pieces and then transferred into a column in the FastDNA Spin Kit (MP Biomedicals). Sodium phosphate buffer (978 µl) and MT buffer (122 µl) were added, and the mixture was homogenized using a Mini-Beadbeater-24 (Biospec Products) at a speed of 3,500 oscl min\(^{-1}\). All subsequent steps for total DNA extraction were carried out according to the manufacturer’s instructions.

The V4 region of the 16S rRNA gene was amplified using the 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GGA CTACHVGGGTWTCTAAAT-3′) primer pair (Caporaso et al., 2011), based on a dual-barcode strategy. For polymerase chain reactions (PCRs), each 50 µl mixture contained 1X PCR buffer, 1.5 mM MgCl\(_2\), 0.2 mM dNTP mix, 0.5 µM of each primer, 2 U Invitrogen Platinum Taq DNA polymerase (Life Technologies) and 2.5 µl DNA extracts from each sample as templates. The PCR program consisted of an initial step at 94°C for 3 min; 30 cycles of 94°C for 45 s, 50°C for 60 s and 72°C for 90 s; and a final extension step at 72°C for 10 min. The PCR products generated were then sequenced using a HiSeq 2500 platform (Illumina) based on a paired-end strategy (2 × 250 bp).

### 2.3 Sequence processing

Raw sequences were processed using the QIIME 2 (v. 2019.1.0) bioinformatics platform (Bolyen et al., 2019). Paired-end sequences were demultiplexed, and primers were trimmed (allowing no error) using cutadapt (Martin, 2011). The sequences were then quality-filtered, merged and dereplicated using the DADA2 workflow (Callahan et al., 2016) to determine amplicon sequence variants (ASVs) representing unique bacterial taxa; no additional trimming or truncation was performed. ASVs appearing in only one sample were discarded. Taxonomy was assigned using the BLAST+ (Camacho et al., 2009) consensus classifier (Bokulich et al., 2018) against the SILVA 132 database (\(p = 10^{-5}\)) (Quast et al., 2013). Eukaryota, archaea, chloroplast, mitochondria and unassigned sequences were removed. Representative sequences were aligned using MAFFT.
(Katoh & Standley, 2013), and a phylogenetic tree was constructed using FastTree (Price et al., 2010). The ASV table was subsampled to 11,572 sequences per sample 100 times (hereafter, bootstraps) for further analyses.

2.4 | Diversity estimation

Faith’s phylogenetic diversity (PD), Shannon diversity and Pielou’s evenness were estimated (based on 100 bootstraps) using the picante (PD) and vegan (Shannon diversity and Pielou’s evenness) packages in R. Differences in the means among the FL, NA and MA fractions were evaluated using one-way ANOVA, and the pairwise groups were further evaluated using Tukey’s HSD test. In addition, abundant ASVs were defined as ASVs accounting for >1% of the relative abundance within any sample (Wu et al., 2017). Ternary plots were used to visualize the preference of abundant ASVs among the FL, NA and MA communities (Smith, 2017), which was indicated by orientation towards a corner using a threshold value of >60%.

Unweighted and weighted UniFrac distances were calculated with the rarefied ASV tables using the GUniFrac package in R (Chen et al., 2012). Principal coordinate analysis (PCoA) was performed on the weighted UniFrac dissimilarity, and ellipses were generated with 95% confidence intervals using the car package (Fox & Weisberg, 2019).

We analysed the distance-decay of community similarity relationship, which represents the decrease in similarities among communities (1 - community dissimilarity) with an increase in geographical distance (Green & Bohannan, 2006). Weighted UniFrac dissimilarity was used because relative abundances are informative for community structuring (Anderson et al., 2011). For samples taken from the surface and bottom layers at the same station, pairwise geographical distances were estimated based on the difference in water depth, which was ignored for estimations at different stations.

Constrained analysis of principal coordinates (CAP) using the vegan package allowed us to relate weighted UniFrac dissimilarities (standardized) with environmental variables. Prior to CAP analyses, forward selection (999 permutations) was run, followed by post hoc fittings using the envfit function (999 permutations) to select environmental variables that significantly accounted for dissimilarities among communities.

2.5 | Null model analysis

To determine whether phylogenetic turnover can be used to infer ecological processes (Stegen et al., 2013), we tested phylogenetic signals using Mantel correlograms as previously described (Dini-Andreote et al., 2015). Briefly, we calculated the relative abundance-weighted means of ASVs for each environmental variable. The resulting mean for each ASV represented its environmental optima (i.e. ecological niches). Differences in environmental optima between ASVs were represented as Euclidean distances using normalized axes. Phylogenetic distances between ASVs were calculated using FastTree with the picante package in R (Kembel et al., 2010). Pearson’s correlation coefficients of differences in environmental optima and phylogenetic distances were quantified based on Mantel correlograms using the vegan package in R (mantel.correlogram function, 50 phylogenetic distance classes and 999 permutations associated with progressive Bonferroni correction).

We analysed the phylogenetic structure of local communities to obtain insights into the drivers of community assembly. Because of the possibility that conservatism in bacterial traits occurs mostly at terminal levels in the phylogeny, we calculated the nearest taxon index (NTI) that quantified the terminal structure of phylogenetic clustering and overdispersion (Webb et al., 2002) using the picante package in R. NTI measures the deviation of observed mean nearest taxon distance (MNTD) from the null expectation (taxa.labels model, 999 randomizations). An NTI of >+2 indicates that the ASVs in a local community are more closely related than expected by chance, suggesting the role of selective pressures (e.g. environmental conditions) in phylogenetic clustering. An NTI of <-2 represents phylogenetic overdispersion, indicating two possible biotic interactions: competition and facilitation. In contrast, a mean NTI across multiple communities that is significantly greater or less than zero indicates phylogenetic clustering or overdispersion, respectively (Zhou & Ning, 2017).

We further assessed the relative importance of heterogeneous selection, dispersal limitation (combined with drift), drift (acting alone), homogenizing dispersal and homogeneous selection using a two-step framework (Stegen et al., 2013). In the first step, we calculated intercommunity phylogenetic turnover using the mean nearest taxon distance (βMNTD). Then, we estimated the β-nearest taxon index (βNTI), representing the difference between the observed βMNTD and the mean null expectation in units of standard deviation (999 randomizations). βNTI >+2 or <-2 signified heterogeneous selection or homogeneous selection, respectively. In the second step, we calculated the Bray-Curtis-based Raup-Crick metric (RCbray) when −2 < βNTI <+2. Homogenizing dispersal, drift and dispersal limitation were inferred based on RCbray values of <-0.95, between −0.95 and +0.95, and >+0.95, respectively. We performed the null model analysis with 100 bootstraps (using the 100 rarefied ASV tables), and the averages (in %) demonstrated the relative importance of different ecological processes.

2.6 | Core ASV analysis

Core microbiomes, which are crucial for community functions (Shade & Handelsman, 2012), are commonly defined based on cut-off values of shared taxa occurrence across communities (Hernandez-Agreda et al., 2017). We defined core ASVs separately in the FL, NA and MA fractions using a 100% threshold (i.e. present in all 38 samples in each group). For phylogenetic analyses, sequences of core ASVs were aligned using MAFFT, and a maximum likelihood tree was constructed using a model recommended by MEGA X (Kumar et al., 2017).
et al., 2018). Moreover, Levin’s habitat niche breadths of core ASVs were calculated using the spa package in R with the following formula (Zhang, 2016):

\[ B_j = \frac{1}{N} \sum_{i=1}^{N} P_{ij} \]

where \( B_j \) represents the habitat niche breadth of ASV \( j \) and \( P_{ij} \) is the proportion of ASV \( j \) in community \( i \) among a total of \( N \) communities.

3 | RESULTS

3.1 | Alpha and beta diversity

Sampling stations exhibited high degrees of heterogeneities with horizontal (inshore and offshore) and vertical (surface and bottom) variations (Figure 1). Factors such as salinity were highly variable, ranging from 0.5 PSU in the surface layer at Station A03 to 34.5 PSU in the bottom layer at Station A16. Hypoxia, indicated by low concentrations of dissolved oxygen (\(<2 \) mg/l), was observed in the bottom layers of five stations (A09, A10, A11, F301 and F404). Salinity, which was significantly (\( p < .05 \)) correlated with most of other variables (13 out of 15), was the most representative factor (Figure 1b). The first axis of the principal component analysis, which indicated inshore-offshore gradients, accounted for 35% of regional variations (Figure S1).

A total of 11,282 ASVs were detected (Figure 2a), which increased in number in the following order: FL (\( n = 3,020 \)) < NA (\( n = 5,283 \)) < MA (\( n = 8,332 \)) communities. A total of 220 abundant ASVs were identified (Figure 2b), of which 109 exhibited no preference among the three types of communities, with < 60% in all types of communities. NA and MA communities had significantly (\( p < .05 \)) higher PD than that of FL communities (Figure 2c), although the Shannon diversity indexes of the three categories did not significantly (\( p > .05 \)) differ from each other (Figure 2d). The MA communities showed significantly lower evenness (\( p < .05 \)) than that of the other two categories (Figure 2e). Gammaproteobacteria (35%), Alphaproteobacteria (19.6%), Cyanobacteria (16.6%) and Bacteroidetes (10.2%) represented the four most abundant groups with relative abundances higher than 10% across all 114 samples (Figure S2).

**Figure 2** (a) Venn diagram showing the number of ASVs among the FL, NA and MA communities. (b) Ternary plot of the distributions of the 220 abundant ASVs (i.e. accounting for >1% of the relative abundance within any sample). The symbol sizes indicate the mean relative abundances of ASVs across all 114 samples, while symbol colours indicate taxonomic classifications. The preference of abundant ASVs among the FL, NA and MA communities is indicated by orientation towards a corner based on a threshold value of >60% (i.e. the grey areas). Box plots showing observations of PD (c), Shannon diversity (d) and Pielou’s evenness (e). In the panels c–e, each circle represents an average of 100 bootstraps, and the different letters (i.e. “a” and “b” in box plots) indicate significant (\( p < .05 \)) differences in means evaluated using one-way ANOVA associated with Tukey’s HSD test.
The FL, NA and MA communities differed significantly but weakly according to analysis of similarity tests based on either unweighted ($R = 0.27, p < .001$) or weighted ($R = 0.18, p < .001$) UniFrac dissimilarities. Thus, the PCoA plots of FL, NA and MA communities largely overlapped (Figure 3), especially in the weighted UniFrac-inferred plots. Moreover, horizontal and vertical turnovers (weighted UniFrac dissimilarities) of the communities were significantly ($p < .05$) correlated with a few factors, such as bottom depth/SiO$_3$ (serving as proxies of freshwater and marine habitats) and sampling depth/temperature (serving as proxies of surface and bottom habitats), respectively (Figure 4). All the three types of communities displayed significant ($p < .0001$) distance-decay relationships (Figure S3).

### 3.2 Core ASVs

A total of 40 core ASVs were identified (Figure 5a), which showed wide niche breadths (minimum = 4.1) (Figure 5b). The core ASVs within the FL, NA and MA fractions accounted for average relative abundances of 36.5%, 24.2% and 26.7%, respectively (Figure 5c), and most of them (34 out of 40) belonged to abundant ASVs (Figure 2b). Based on phylogenetic analyses, a total of 13 core ASVs were closely related to hydrocarbonoclastic bacteria (Figure 5d).

### 3.3 Community assembly

Significant phylogenetic signals ($p < .01$) were observed across relatively short phylogenetic distances (<20% of the maximum phylogenetic distance) (Figure S4), enabling the use of phylogenetic turnover to infer ecological processes. The phylogenetic structure of most (FL, $n = 38$; NA, $n = 34$; MA, $n = 36$) local communities tended to cluster according to the estimates (i.e., >+2) of the NTI (Figure 6a). Null model analyses showed that although the relative importance of homogeneous selection was predominant, it decreased in the following order: FL (90%) > NA (86.8%) > MA (73.3%) communities (Figure 6b). In contrast, there was an increase in dispersal limitation according to the following order: FL (8.6%) < NA (11.7%) < MA (13.2%). The importance of drift was prominent only in MA communities (11.2%), while those of both heterogeneous selection and homogenizing dispersal were minor (e.g. heterogeneous selection in the MA fraction was <1.8%) in all three types of communities.

### 4 DISCUSSION

#### 4.1 Domination of homogeneous selection

Our results suggested that homogeneous selection could be the most important ecological process in shaping bacterial communities across the FL, NA and MA fractions (Figure 6b). This scenario is somewhat unexpected in the perturbed SCSPR, characterized by marked environmental heterogeneities (Figure 1), since highly variable factors such as salinity can heterogeneously select coastal bacterial communities (Campbell & Kirchman, 2013). In contrast, homogeneous selection results in low levels of variation in communities, which are imposed by low environmental variability. However, the shallow distance-decay relationships (Figure S3) support the dominance of homogeneous selection, considering that homogeneous selection contributes to high community similarities (Dini-Andreote et al., 2015). The dominance of homogeneous selection is also consistent with the findings of a recent study showing that homogeneous selection dominated the assemblies of both FL and particle-associated bacterial communities in the inner bay of the SCSPR (Wang et al., 2020).

Thus, we propose that biogeochemical processes, especially extensive eutrophication, may result in a relatively homogeneous environment for bacterial community structuring in the SCSPR. In the SCSPR, eutrophication due to the influx of large volumes of terrestrial nutrients, which causes phytoplankton blooms and the decomposition of organic particles, is primarily responsible for the formation of a hypoxic layer in the summer months (Lu et al., 2018). Consistent with eutrophic conditions, we detected high concentrations of surface chlorophyll a (maximum, 11.2 µg\text{\textit{l}}^{-1} at Station F404) and low concentrations of dissolved oxygen (minimum, 1 mg\text{\textit{l}}^{-1} at Station F404) in the bottom layer. Cyanobacteria, mostly represented by Synechococcus, were dominant in the communities found at many of the surface stations, with a relative abundance of 28.5% (see Figure S2a). Synechococcus exhibited bloom-like cell concentrations, with a maximum of $6.5 \times 10^5$ cells

![Figure 3](https://example.com/figure3.png) Principal coordinate analysis (PCoA) plots of (a) unweighted and (b) weighted UniFrac distances between bacterial communities. FL, NA and MA samples are represented by red, blue and green symbols (dots, surface: triangles, bottom), respectively, in ellipses and are associated with 95% confidence intervals.
and an average of $1.5 \times 10^5 \text{ cells ml}^{-1}$ in the surface samples, similar to previous *Synechococcus* blooms in the SCSPR, which were frequently observed during the wet season (Li et al., 2019). *Synechococcus* blooms in the SCSPR can be an indicator of eutrophic status (Rajaneesh et al., 2015). Bacterial assemblages besides *Synechococcus* might largely be shaped through the recruitment of taxa related to *Synechococcus* blooms (Buchan et al., 2014; Needham & Fuhrman, 2016). The bacteria–phytoplankton interaction acts as a relatively independent driver of bacterial communities when compared with physiochemical factors such as salinity (0.5–34.5 PSU in this study) (Kirchman et al., 2017). Consequently, *Synechococcus* and pigmented eukaryotes are disproportionately important, considering their key roles in regulating FL, NA and MA communities (Figure 4) versus their relatively minor contributions to environmental heterogeneities (Figures 1b and S1).

Moreover, eutrophication levels of coastal ecosystems can adversely affect the phylogenetic structure of local communities (Dai et al., 2017). Correspondingly, most phylogenetic structure of the local communities in this study were closely clustered (Figure 6a), indicating that the bacterial communities have undergone heavy selection (Zhou & Ning, 2017). This phylogenetic clustering is, therefore, consistent with the dominance of homogeneous selection at the regional scale (Figure 6b). In line with our finding, recent studies have reported temporal dominance of homogeneous selection during phytoplankton blooms in a eutrophic coast (Zhang et al., 2018) and a eutrophic urban lake (Xu, Zhao et al., 2020).

We further propose that eutrophic conditions can impose homogeneous selection on potentially native bacterial communities in the SCSPR. First, we suggest the existence of native bacterial communities in the SCSPR rather than a simple mixture of freshwater and seawater communities (Crump et al., 2004). The productive SCSPR harbours bacterial communities with high growth rates; however, these bacterial communities have longer residence periods in the SCSPR compared to riverine systems due to the largely reduced level of directional transportation. These two conditions facilitate the formation of native coastal bacterial communities (Crump et al., 2004). Notably, native bacterial communities can adapt to copiotrophic environments, even those with long-term disturbances (Xiong et al., 2015), through a combination of genomic architecture (Lauro et al., 2009) and metabolic strategies such as nutrient costs (Grzymski & Dussaq, 2012), leading to a consistent change in the compositions of communities across large gradients (Fodelianakis et al., 2014). Thus, homogeneous selection can be enhanced in potentially native bacterial communities because wide adoptions to eutrophic conditions strengthen community structures rather than generate variations.

### 4.2 Community assembly across multiple fractions

Free-living, NA and MA bacterial communities share remarkable similarities in their community assembly patterns, as demonstrated by the dominance of homogeneous selection (Figure 6b). Suspended particles in coastal ecosystems are usually old (i.e. mineral-rich) and mainly comprise inorganic matter of terrestrial and sediment origins (Dang & Lovell, 2016). The nature of these particles may...
largely minimize the effect of particle colonization, which results in different assemblages among FL, NA and MA communities (Dang & Lovell, 2016), unlike organic content-rich particles in pelagic waters with largely differentiated FL and particle-associated bacterial communities (Ortega-Retuerta et al., 2013). Moreover, it has been reported that FL and particle-associated bacterial communities likely interact with each other through their attachment to and detachment from diverse particles (Kierboe et al., 2003). Such interactions may homogenize the variations among the three kinds of bacterial communities, resulting in many (n = 1,647) shared ASVs (Figure 2a). Similarly, both unweighted and weighted UniFrac-based PCoA plots of FL, NA and MA communities largely overlapped (Figure 3), especially in the case of the latter, indicating that the abundance of ASVs dampens the differences among the three categories; this is consistent with the finding that approximately half of the abundant ASVs show no preference among the FL, NA and MA communities (Figure 2b).

However, the pattern of community assembly was also considerably different across the different fractions (Figure 6b). First, dispersal limitation among the FL, NA and MA communities is increased, likely because water movement-driven passive dispersal is less limited in FL bacteria compared to that in particle-associated bacteria (Villarino et al., 2018). Second, the chemical properties of particles are diverse (Zhang et al., 2016), and thus, particles have higher substrate availabilities (i.e. niches) in their microenvironments than in the FL habitat (Mestre et al., 2017), resulting in significantly higher PD estimates in NA and MA than in FL communities (Figure 2c) (Xu, Zhao et al., 2020). At the same time, locally enriched substrates can largely increase the relative abundance of favoured groups, leading to uneven community structures. Therefore, the three fractions showed non-significant differences in their proportional Shannon diversities (Figure 2d), despite the significantly higher PD estimates of NA and MA communities. In particular, MA communities exhibited significantly lower evenness than both FL and NA (Figure 2e). The large differences in PD and evenness among MA communities indicate the occasional presence/absence of many taxa in some microparticles due to distinct substrate availabilities, which tends to trigger drift (Figure 6b) (Xun et al., 2019). Taken together, deterministic (mainly represented by homogeneous selection) and stochastic (mainly represented by dispersal limitation and drift) processes decrease and increase, respectively, across the FL, NA and MA fractions.

4.3 | Core ASVs

The core ASVs identified in this study indicate that anthropogenic activities may markedly influence the assembly of bacterial communities in the SCSPR (Figure 5). In particular, the most abundant core ASV01 shares 100% similarity with the strain Acinetobacter venericolus BU71, isolated from the surface water of petroleum-contaminated coastal sites (Figure 5d) (Mahjoubi et al., 2013). There were nine core ASVs belonging to hydrocarbonoclastic bacteria or taxa with cooperative interactions, which are closely related to taxa detected in an oil sheen in the Gulf of Mexico (McGenity et al., 2012). Obligate hydrocarbonoclastic bacteria have long been recognized for their roles in the removal of petroleum hydrocarbons from polluted coastal waters (Yakimov et al., 2007). The influx of hydrocarbons has been a major concern in the SCSPR, both in the waters (Wang et al., 2007) and sediments (Chen et al., 2006), especially during the wet season (Wang et al., 2007). Crude oil represents the most complex mixture of organic compounds on Earth (Head et al., 2006), and thus, hydrocarbonoclastic bacteria have evolved a versatile range of metabolic functions (Mason et al., 2012). Hydrocarbonoclastic bacteria are widely found in the SCSPR (Wu et al., 2014), suggesting that the SCSPR is environmentally favourable for these bacteria.

Importantly, the core ASVs accounted for large proportions of the whole communities (Figure 5c,d). For example, the most abundant core ASV01 had remarkable relative abundances (FL, 6%; NA, 6.3%; MA, 9.1%) (Figure 5d), indicating that core ASVs can influence responses to coastal environments. Notably, high relative abundances and widespread occurrence of core ASVs reduce dissimilarities among communities and contribute to the relative importance of homogeneous selection as a primary mediator of community assembly (Ostman et al., 2010). According to the threshold value (i.e. >3) of habitat niche breadths (Logares et al., 2013) (Figure 5b), all core ASVs are habitat generalists; this is consistent with that they contribute to the dominance of homogeneous selection owing to their widespread occurrence. Our findings support the notion that bacterial communities are indicators of contaminant stress (e.g. due to hydrocarbon pollutants) in coastal zones (Duran & Cravo-Laura, 2016; Sun et al., 2012).

Finally, we must point out that the bacterial communities observed in this study represent a temporal snapshot of the SCSPR. Coastal environments are highly dynamic, and fine-scale temporal variability shows the turnover time of bacterial communities to be...
just a few days (Bunse & Pinhassi, 2017; Osterholz et al., 2018). The bacterial community patterns observed in this study have rarely been reported in the SCSPR, despite numerous previous studies being conducted there. Considering hydrocarbонolytic bacteria as an example, they might be highly diverse (McKew et al., 2007), although serving as a minor component of the native bacterial communities in the SCSPR. The ability of hydrocarbонolytic bacteria to thrive depends on the prevailing physicochemical factors in particular (Yakimov et al., 2007). Since our samples were collected immediately before the tropical storm Roke (July 21–23, 2017), we speculate that a combination of pre-storm conditions (e.g. low wind speed and intense sunlight) might have promoted the occurrence of oil-degrading bacteria (Bacosa et al., 2015) and contributed to the formation of temporally unique bacterial communities in the SCSPR (Bryant et al., 2016). Therefore, we suggest that more efforts are urgently needed to adequately understand the mechanism underlying the assembly of bacterial communities in complex coasts.

5 | CONCLUSIONS

In this study, the observed ASVs were most abundant in the FL, followed by the NA and MA communities, in the SCSPR. Approximately half of the abundant ASVs showed no preference among the three fractions. The detection of a few core ASVs, which are closely related to oil-degrading bacteria, indicates the effect of anthropogenic activities on bacterial communities. The FL, NA and MA communities differed significantly but weakly. Importantly, homogeneous selection might have been the primary factor shaping the bacterial communities during the sampling time, although more investigations are needed to achieve a better understanding of bacterial community assembly in the SCSPR. Moreover, subdividing the whole community into FL, NA and MA fractions is helpful for revealing new information about their underlying assembly mechanisms; we observed that deterministic processes (mainly represented by homogeneous selection) were most predominant in FL, followed by NA and MA communities, whereas stochastic processes (mainly represented by dispersal limitation and drift) exhibited the reverse order. In summary, this study provides an unexpected snapshot of bacterial communities in the SCSPR, filling knowledge gaps regarding bacterial community assembly. The results of this study may inspire future studies focused on anthropogenic influences on coastal ecosystems.

ACKNOWLEDGEMENTS

This work was supported by the Theme-based Research Scheme of the Hong Kong Research Grants Council (No. T21/602/16), the Fundamental Research Funds for the Central Universities (No. 19lgpy89) and the Regional Marine Geological Survey Projects of China Geological Survey (No. DD20191002). We acknowledge the support of SML311019006/311020006. We thank P. Y. Lee for the flow cytometry analysis.

[Correction added on 18 November 2020, after first online publication: the Acknowledgements have been updated.]

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13193.

DATA AVAILABILITY STATEMENT

Raw sequences have been deposited in the Sequence Read Archive (NCBI) under the accession number PRJNA634806.

ORCID

Wenxue Wu https://orcid.org/0000-0003-2196-2594

REFERENCES

Aguirre, M., Abad, D., Albaina, A., Cralle, L., Goñi-Urriza, M. S., Estonba, A., & Zarraona-Jaíndia, I. (2017). Unraveling the environmental and anthropogenic drivers of bacterial community changes in the Estuary of Bilbao and its tributaries. *PLoS One*, 12(6), e0178755. https://doi.org/10.1371/journal.pone.0178755

Andersen, K. H., Berge, T., Gonçalves, R. J., Hartvig, M., Heuschele, J., Hylander, S., Jacobsen, N. S., Lindemann, C., Martens, E. A., Neuheimer, A. B., Olsson, K., Palacz, A., Prowe, A., Sainmont, J., Traving, S. J., Visser, A. W., Wadhwa, N., & Kierboe, T. (2016). Characteristic sizes of life in the oceans, from bacteria to whales. *Annual Review of Marine Science*, 8(1), 217–241. https://doi.org/10.1146/annurev-marine-122414-034144

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., Sanders, N. J., Cornell, H. V., Comita, L. S., Davies, K. F., Harrison, S. P., Kraft, N. J. B., Stegen, J. C., & Swenson, N. G. (2011). Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. *Ecology Letters*, 14(1), 19–28. https://doi.org/10.1111/j.1461-0248.2010.01552.x

Arnosti, C. (2011). Microbial extracellular enzymes and the marine carbon cycle. *Annual Review of Marine Science*, 3(1), 401–425. https://doi.org/10.1146/annurev-marine-120709-142731

Azam, F., & Long, R. A. (2001). Sea snow microcosms. *Nature*, 414(6863), 495–498. https://doi.org/10.1038/35107174

Bacosa, H. P., Liu, Z., & Erdner, D. L. (2015). Natural sunlight shapes crude oil-degrading bacterial communities in northern Gulf of Mexico surface waters. *Frontiers in Microbiology*, 6, 1325. https://doi.org/10.3389/fmicb.2015.01325

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., Hutten, G. A., & Gregory Caporaso, J. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2’s q2-feature-classifier plugin. *Microbiome*, 6(1), 90. https://doi.org/10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghali, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Carballa-Rodriguez, A. M., Chase, J., ... Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, 37(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bryant, J. A., Aylward, F. O., Epplle, J. M., Karl, D. M., Church, M. J., & DeLong, E. F. (2016). Wind and sunlight shape microbial diversity in surface waters of the North Pacific Subtropical Gyre. The *ISME Journal*, 10(6), 1308–1322. https://doi.org/10.1038/isme.2015.221

Buchan, A., LeCleir, G. R., Gulvik, C. A., & Gonzalez, J. M. (2014). Master recyclers: Features and functions of bacteria associated with
BIOSKETCHES

Wenxue Wu is an Associate Professor in the School of Marine Sciences at the Sun Yat-sen University. His research interests mainly focus on marine microbial diversity and community assembly.

Hongbin Liu is a Professor in the Department of Ocean Science at The Hong Kong University of Science and Technology. His research interests mainly focus on factors that control microbial biomass, diversity, production and metabolism in different marine ecosystems.

Author contributions: The study was conceived by W.W. and H.L.; Z.X. collected the field samples; W.W. performed the molecular laboratory work and sequencing data analyses; M.D. and J.G. contributed the environmental data; W.W. wrote the first draft of the manuscript; and all authors contributed substantially to revisions.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Wu W, Xu Z, Dai M, Gan J, Liu H. Homogeneous selection shapes free-living and particle-associated bacterial communities in subtropical coastal waters. Divers Distrib. 2021;27:1904–1917. https://doi.org/10.1111/ddi.13193