MLST

REVIEW

Microbial assembly, interaction, functioning, activity and diversification: a review derived from community compositional data

Jiwen Liu[1,2,3] · Zhe Meng[1] · Xiaoyue Liu[1] · Xiao-Hua Zhang[1,2,3]

Abstract

Microorganisms play crucial roles in maintaining ecosystem stability. The last two decades have witnessed an upsurge in studies on marine microbial community composition using high-throughput sequencing methods. Extensive mining of the compositional data has provided exciting new insights into marine microbial ecology from a number of perspectives. Both deterministic and stochastic processes contribute to microbial community assembly but their relative importance in structuring subcommunities, that are categorized by traits such as abundance, functional type and activity, differs. Through correlation-based network analysis, significant progress has been made in unraveling microbial co-occurrence patterns and dynamics in response to environmental changes. Prediction of ecosystem functioning, based on microbial data, is receiving increasing attention, as closely related microbes often share similar ecological traits and microbial diversity often exhibits significant correlations to ecosystem functioning. The ecosystem functioning is likely executed not by the whole community, but rather by an active fraction of a community, which can be inferred from the marker gene transcription level of community members. Furthermore, the huge amount of microbial community data has significantly expanded the tree of life and illuminated microbial phylogenetic divergence and evolutionary history. This review summarizes important findings in microbial assembly, interaction, functioning, activity and diversification, highlighting the interacting roles of different aspects, derived from community compositional data.

Keywords rRNA gene · Microbial community · Assembly · Interaction · Ecosystem functioning

Introduction

Microorganisms play key roles in biogeochemical cycling that are fundamental in maintaining climate and ecosystem stability. The structure of microbial communities is closely associated with environmental conditions and therefore is likely to evolve in the context of global change (Gutknecht et al. 2012). In the marine environment, frequent natural events and increasing human activity dramatically influence microbial community dynamics, which will change the balance of biogeochemical cycles and alter ecosystem functioning (Hutchins and Fu 2017). One of the major concerns associated with global changes is how to effectively predict variations in ecosystem functioning. Microorganisms, as major drivers of many biogeochemical processes, provide a linkage between ecosystem functioning and environments (Singh et al. 2010).

Marine microbial communities are significantly affected by environmental changes. Sanger and high-throughput sequencing in recent decades have provided an enormous amount of sequence data of molecular marker genes, including the ribosomal RNA (rRNA) gene. These data have helped to provide insights into marine microbial community dynamics (Liu et al. 2019; Needham et al. 2017; Reji et al. 2019), which are driven by environmental factors, such as salinity

✉ Xiao-Hua Zhang
xhzhang@ouc.edu.cn

[1] MOE Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, Qingdao 266003, China

[2] Laboratory for Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China

[3] Institute of Evolution and Marine Biodiversity, Ocean University of China, Qingdao 266003, China

(Lozupone and Knight 2007) and temperature (Sunagawa et al. 2015). However, the association between environments and microbial dynamics (termed as deterministic processes) can be confounded by the effects of random events (termed as stochastic processes that include ecological drift and dispersal) (Mo et al. 2018; Wang et al. 2019; Zhou and Ning 2017). Deterministic and stochastic processes, which jointly determine microbial biogeography, vary in their relative contribution to community assembly over different temporal and spatial scales (Zhou and Ning 2017). Differences in microbial distribution impact ecosystem functioning. Their relationships can mostly be explained by the observed correlations between microbial phylogeny and functional traits and between microbial diversity and ecosystem functioning. There are ways to predict microbial functional potential based on taxonomy (Aßhauer et al. 2015; Langille et al. 2013; Louca et al. 2016). Therefore, it has become increasingly common to improve the fitness of ecosystem functioning prediction by including microbial data.

Marine microbes are highly diverse and encompass taxonomically and functionally different lineages. Complex interactions occur among microbial taxa, which underpin community stability and functioning. However, elucidation of microbial interactions is challenging and is largely dependent on correlation-based network analysis (Liu et al. 2014b; Milici et al. 2016; Zhang et al. 2014; Zhou et al. 2018). Diverse microbial communities can be divided into subcommunities, based on different criteria, such as abundance (abundant and rare taxa), functional type (e.g., autotrophic and heterotrophic taxa) and activity (active and dormant taxa). Accumulating evidence is showing that microbial subcommunities differ in their environmental sensitivity, interaction and distribution patterns (Wu et al. 2017; Zhang et al. 2014). Thus, different subcommunities may represent different consortiums and differ in their roles in ecosystem functioning prediction. Currently, the ecology of microbial subcommunities is less understood than the whole community, raising the need for a resolved community-based classification for future analysis.

By implementing rRNA gene-based amplicon sequencing, numerous previously unknown microbial lineages, even at the phylum level, have been described from the marine environment (Brown et al. 2015; DeLong 1992; Inagaki et al. 2003). These, together with those identified from terrestrial habitats, dramatically expand the tree of life (Hug et al. 2016). Refined phylogenetic analysis of molecular marker genes further demonstrates the occurrence of habitat-specific ecotypes within a lineage (Ivars-Martinez et al. 2008; Liu et al. 2014a). The diversification of microorganisms can be attributed to a joint effect of genetic and environmental variabilities, which dictate the specific evolutionary history of a taxon.

In this review, five aspects are presented (microbial assembly, interaction, functioning, activity and

diversification) to show how microbial community data contribute to the understanding of marine microbial ecology. They are organized along a stepwise understanding of the role microbial communities play in marine ecosystems. A synthetic view of these aspects will provide novel insights into their interactions and complementarity, which will in turn help to stimulate new ideas on the interpretation of community compositional data and the perception of new microbial community studies.

Processes of microbial community assembly

One long debated question in community ecology is which processes determine an ecological community to assemble (Preston 1948). The current, well-established theories are primarily derived from research on animals and plants. Studies on microorganisms are scarce because of the assumption that microorganisms do not have distribution patterns due to their large numbers and small size (Baas-Becking 1934). However, using advanced sequencing technologies and statistical methods, microbial distribution patterns have been discovered in many natural environments including seawater and marine sediments (Liu et al. 2014b, 2015a; Lozupone and Knight 2007; Martiny et al. 2006). The spatial turnover of microbial communities always reflects a distance-decay relationship and/or a taxa-area relationship, which are the two most well established patterns depicting increasing community dissimilarity with spatial distance (Nekola and White 1999) and increasing taxa richness with area size (Horner-Devine et al. 2004), respectively.

A framework has been established that considers the niche and neutral theories as potential mechanisms underpinning microbial biogeography (Dini-Andreote et al. 2015; Dumbrell et al. 2010; Leibold and McPeek 2006; Table 1). The long standing niche theory shows how biodiversity is structured by physiochemical (environmental heterogeneity) and biotic (inter taxa interaction) factors (Chase and Leibold 2003). In niche theory, every taxon is assumed to have unique and non-overlapping traits, which enable it to exert different responses/effects on an environment and to occupy different ecological niches (Leibold 1995). This heterogeneity in traits also eliminates inter taxa competition by resource partition and thus allows an infinite number of taxa to coexist in an environment, leading to high microbial diversity (Leibold and McPeek 2006). In this context, niche theory cannot explain the observed differences in the abundance of different taxa, although microbial abundance under niche theory always follows a log-normal distribution pattern (Hubbell 2001). Later, Hubbell (2001) proposed a neutral ecological view on the assembly of taxa-rich communities that consisted of many rare taxa. Unlike niche theory, the neutral theory assumes that all individuals are ecologically

Table 1 Comparison of the niche and neutral theories

|	The niche theory	The neutral theory
Time proposed	1959 (A general description of niche concept)	2001
Concept	Coexistence by niche differentiation	Random events occurring on individual taxon
Traits	Functionally different	Functionally equivalent
Relative abundance	Log-normal distribution	Zero-sum multinomial distribution
Processes	Abiotic environmental selection and biotic interspecific interactions	Ecological drift and dispersal
Feature of process	Deterministic	Stochastic

equivalent, share the same way of life and have a similar response/effect on environments. Therefore, interactions among microbes, and between microbes and environments are ignored in the neutral theory (Hubbell 2001). Microbial diversity is controlled by randomly occurring events, such as ecological drift (change in the relative abundance of taxa in a location due to chance demographic fluctuations) and dispersal (movement of taxa across spaces) (Chave 2004; Hanson et al. 2012; Hubbell 2001). Because such ecological events happen differently to each individual taxon, the relative abundance of different taxa can be partially explained by the neutral theory (Volkov et al. 2003).

Microbial ecology studies have consistently considered abiotic selection as a niche-based (deterministic) process (also known as habitat filters or species sorting) and ecological drift and dispersal limitation as neutral-based (stochastic) processes (Fig. 1). Selections are mostly represented by environmental correlations. Drift is difficult to measure and it often interacts with restricted dispersal (dispersal limitation) to produce a distance-decay relationship (Hanson et al. 2012). Dispersal limitation is often represented by spatial distance. However, spatial distance can be a deterministic factor when it correlates to physiochemical factors and both population size (neutral) and traits (niche) can affect the dispersal ability of different taxa (Hanson et al. 2012). These make the relationships between selection, dispersal and stochasticity complex (Evans et al. 2017). To accurately evaluate stochasticity, neutral and null models have been increasingly implemented (Chase and Myers 2011; Jeraldo et al. 2012, Table 2). Nevertheless, these models cannot distinguish different aspects of the stochastic process. To solve this problem, Stegen et al. (2013, 2015) proposed a framework to quantify the relative roles of deterministic and stochastic processes. This framework integrates both the null model and phylogenetic information and assumes that phylogenetically close taxa tend to have similar ecological traits. It contains two steps with the first to estimate the role of selection using phylogenetic dissimilarities and the second to differentiate the role of dispersal and drift using the Bray–Curtis dissimilarities among microbial communities (Jia et al. 2018; Logares et al. 2018; Stegen et al. 2013, 2015; Zhou and Ning 2017).

Deterministic and stochastic processes jointly govern the assembly of microbial communities (Chave 2004). However, their relative importance varies across different spatial and temporal scales (Table 2), depending on the strength of environmental gradients and the sensitivity of the microbes to environmental changes. If the extent of environmental variation is greater than the threshold a microbe can endure, dispersal will be prevented (Wang et al. 2013), leading to the predominance of determinism. Thus, the mechanisms underlying microbial community assembly would alter over a seasonal or longer term period with changes in the magnitude of environmental heterogeneity (Dini-Andreote et al. 2015; Langenheder et al. 2012). These conclusions are mostly based on investigations of terrestrial microbial communities. By comparison, there have been very few studies on the relative roles of deterministic and stochastic processes in the marine environment. However, there is a general perception that stochastic processes have a greater effect on the assembly of planktonic bacterial and archaeal communities than deterministic processes (Table 2). This can be explained by marine prokaryotes having evolved strong adaptation capabilities to environmental changes and by spatial connectivity and seawater movement homogenizing environmental conditions. Another explanation is that the environmental factors analyzed to represent deterministic processes may be not the most relevant ones affecting community variations. Further studies are needed to confirm such a hypothesis and to compare the assembling processes between different habitats, such as coastal water *vs* open ocean and water *vs* sediment.

Different types of marine organisms differ in their responses to deterministic and stochastic processes. Wu et al. (2018) reported that determinism had a stronger effect on planktonic protist communities than on bacterial communities, which may relate to differences in their environmental sensitivity. Such different responses between bacteria and micro-eukaryotes have also been observed in soil (Powell et al. 2015a) and freshwater (Logares et al. 2018) habitats. Additionally, subcommunities that are divided by abundance, activity, functional trait or occupancy, can also undergo different ecological processes (Fig. 1). A microbial community is usually made up of a

Fig. 1 The assembly of microbial communities is governed by a joint effect of deterministic (niche) and stochastic (neutral) processes. Subdivision of a microbial community (according to abundance, activity, function and occupancy) into subcommunities will facilitate an accurate view of microbial assembling processes
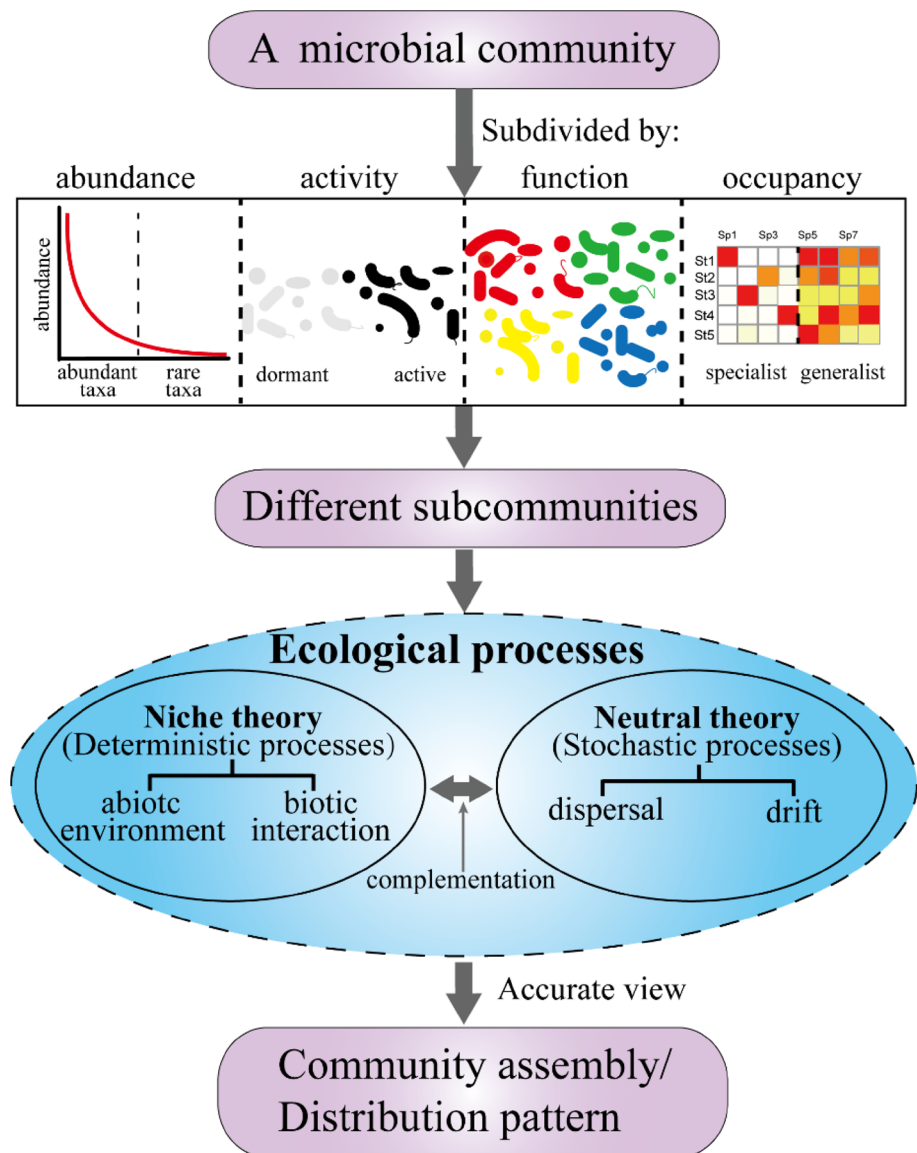
few abundant taxa and a long tail of rarer ones (Pedrós-Alió 2006, 2012). The rare taxa account for a great proportion of the microbial diversity and have been shown to assemble non-randomly and display similar distribution patterns to the abundant taxa (Galand et al. 2009; Gong et al. 2015; Liu et al. 2015b; Mo et al. 2018). Nevertheless, the abundant and rare taxa have both been observed to be differently affected by stochastic and deterministic processes (Liu et al. 2015b; Mo et al. 2018). Mo et al. (2018) reported that the rare bacterioplankton in coastal seawater had a weaker response to environmental factors than abundant taxa; this may be due to the small population size of the rare taxa, making them more susceptible to ecological drift (Nemergut et al. 2013). By contrast, a survey of bacterial communities in freshwater lakes and reservoirs revealed a greater influence of environmental changes on the rare than the abundant taxa (Liu et al.

2015b). These findings suggest complicated microbial ecological responses across distinct ecosystems. In this context, further studies are needed to gain an insight into the assembling processes of abundant and rare taxa in different environments. The most urgent need is to propose a common definition for rare taxa, facilitating a parallel comparison across studies (Jia et al. 2018). Subcommunities divided by functional traits, activity and occupancy receive less attention and have mainly been analyzed in terrestrial environments. For example, in deserts, the phototrophic community was mainly affected by stochastic processes, whereas the heterotrophic community displayed patterns mainly driven by environmental stresses (Caruso et al. 2011). The assembly of generalists and specialists in plateau lakes, however, was driven by stochastic and deterministic processes, respectively (Liao et al. 2016). While these studies provide novel and accurate information about

Table 2 Examples of studies examining the relative role of deterministic and stochastic processes in structuring microbial community

Community	Ecosystem	Spatial scale	Stochasticity measurement	Relative importance		References
				Determinism	Stochasticity	
Terrestrial ecosystem						
Arbuscular mycorrhizal fungi	Soil	< 50 m	Species relative abundance and spatial distance	√	–	Dumbrell et al. (2010)
Bacteria	Desert	Global	Neutral model	Heterotrophic community	Photosynthetic community	Caruso et al. (2011)
Bacteria	Freshwater lakes	~130 km in maximum	Neutral community model	–	√	Roguet et al. (2015)
Bacteria	Soil spanning 105 years succession	–	Newly designed model	Community when succession proceeded	Community at initial stage	Dini-Andreote et al. (2015)
Bacteria and archaea	Groundwater	–	Stegen's framework	√	–	Graham et al. (2017)
Marine ecosystem						
Microbial eukaryotes	Coastal water and sediment	~20 km in maximum	Neutral community model and variation partitioning		√	Chen et al. (2017)
Bacteria	Coastal water	~200 km	Stegen's framework	High eutrophication	Low eutrophication	Dai et al. (2017)
Bacteria and protist	Seawater	51–826 km	Stegen's framework and variation partitioning	Protist community (bottom water)	Protist (surface and deep chlorophyll maximum waters) and bacterial (all three layers) communities	Wu et al. (2018)
Bacteria	Coastal water	~1300 km in maximum	Neutral model and variation partitioning	–	Both abundant and rare communities	Mo et al. (2018)
Archaea	Coastal water	~200 km	Stegen's framework and variation partitioning	–	√	Wang et al. (2019)

the distribution patterns and assembling processes of microbial communities, there is an urgent need to investigate different subcommunities in the marine environment. It should be noticed that when evaluating the relative role of determinism and stochasticity, the estimated contribution from determinism is largely affected by the set of environmental factors measured, since they are not necessarily the most relevant parameters that provide the best explanatory power for the community variations.

Patterns of microbial co-occurrence

In niche theory, the microbe-microbe interactions, although being ecologically important, are less understood compared to the microbe-environment relationships (Chase and Leibold 2003). Inclusion of interactions to explain microbial distribution patterns is a great challenge, largely due to the difficulty in obtaining microbial co-cultures. An alternative way of elucidating microbial interactions is to apply correlation-based network analysis (Barberán et al. 2012; Layeghifard et al. 2017; Weiss et al. 2016), which is enhanced by the increase of community compositional data and the development of statistical tools. The most popular method used for constructing a correlation-based network is to calculate the Spearman's rank correlation coefficients between taxa (Barberán et al. 2012; Table 2). Other methods are also available, including SPIEC-EASI, CCLasso, REBACCA, CoNet, SparCC, WGCNA, Molecular Ecological Networks Analysis, Local Similarity Analysis, Maximal Information Coefficient, etc. These methods, however, vary in their sensitivity and precision (Layeghifard et al. 2017; Weiss et al. 2016).

Nodes and edges are fundamental components of a network, representing taxa and correlations, respectively. Edge thickness often denotes the degree of a correlation, with a

thicker edge representing a higher correlation coefficient. On the basis of nodes and edges, a number of parameters can be calculated to represent the topological structure of a network, including degree, density, betweenness centrality, network diameter and clustering coefficient (Newman 2003). The degree of a node describes its connectivity to other nodes, with a higher value indicating a wider correlation. The betweenness centrality of a node describes the number of shortest paths between any two nodes going through it. The nodes with high degree and low betweenness centrality potentially represent the keystone taxa of a community (Berry and Widder 2014; Liang et al. 2016). The keystone taxa are the cornerstone and initial components for a community to assemble (Berry and Widder 2014) and have recently been defined as "highly connected taxa that individually or in a guild exert a considerable influence on microbiome structure and functioning irrespective of their abundance across space and time" (Banerjee et al. 2018). A group of densely connected nodes with weak correlations to other nodes forms a module. Modular analysis can help to simplify the processes of identifying keystone taxa and/ or exploring the effect of environmental factors on microbe-microbe interactions.

Despite having the potential to infer mutualistic (positive) and antagonistic (negative) effects, the co-occurrence patterns illustrated by the network analysis can have different meanings, such as similar environmental preference and lifestyle, resource partitioning and nutrient cross-feeding that do not involve direct interactions. Network analyses have consistently revealed patterns dominated by positive co-occurrences in both terrestrial and marine bacterial communities (Barberán et al. 2012; Ju et al. 2014; Liu et al. 2014b; Ma et al. 2016; Milici et al. 2016; Zhang et al. 2014; Zhou et al. 2018). This suggests a ubiquitous non hostile process in bacterial community assembling irrespective of habitat. In fact, a network analysis of prokaryotic communities in global surface seawater samples (*Tara* oceans) showed that positive co-occurrences accounted for up to 90% of all correlations (Lima-Mendez et al. 2015; Milici et al. 2016). This may be explained by the similar environmental preference (possibly driven by some unknown/unmeasured environmental factors) and/or high resistance of marine prokaryotes to environmental stresses, which enable them to coexist in the same ecological niche. Indeed, the finding is in line with the observation of a relatively weaker effect of determinism relative to stochasticity on the distribution of marine planktonic prokaryotes as stated above. It is also likely that auxotrophy in marine microbes contributes to the observed positive correlations, since several bacterial groups (such as the SAR11 clade of *Alphaproteobacteria*, Tripp et al. 2009) have been observed to gain fitness by obtaining a biomolecule from other groups. By contrast, a study of planktonic bacterial communities in the South China Sea showed that negative correlations dominated the co-occurrence patterns in active bacteria compared to positive correlations in total bacteria (Zhang et al. 2014). Resource competition may occur in different active bacterial groups. It is noticeable that microbial networks vary with space and time (Table 3). Significant changes in network topological structures have been observed in seawater at different depths (Chow et al. 2013) and between different seasons (Chafee et al. 2018). Additionally, Milici et al. (2016) reported that free-living bacterioplankton possessed highly interconnected networks compared to particle-attached communities. They postulated that the distinct nutrient-utilizing strategies of these two groups might be responsible for such a discrepancy. However, it is unexpected to find more between-taxa connections in the free-living community, since particle-attached microbes have a higher cell abundance and are physically closer to one another. This provides further evidence that the network-derived co-occurrence patterns are not always good proxies of true interactions between microbes.

Several studies have attempted to use networks to infer potential functional couplings between microbes. For example, *Thaumarchaeota* Marine Group I (MG-I), the most abundant archaeal clade in the marine environment, capable of ammonia oxidization, has been found to co-occur with *Nitrospina* (Reji et al. 2019) and/or with *Nitrospira* when *Nitrospina* is absent or in low abundance (Wang et al. 2019), both of which are nitrite oxidizers. Their co-occurrence in seawater is supported by substrate feeding (nitrite produced by MG-I is the substrate of *Nitrospina/Nitrospira*) and facilitates the complete nitrification process. Previous efforts to explore co-occurrence patterns between functional bacteria in marine sediments have demonstrated significant correlations between sulfate-reducing bacteria and sulfur-oxidizing bacteria, and between sulfate-reducing bacteria and nitrite-oxidizing bacteria (Liu et al. 2014b). Elucidation of co-occurrence patterns with functional gene abundance derived from GeoChip and metagenome may facilitate a more direct inference. However, the obtained co-occurrence patterns should be treated with caution when used to infer functional couplings, since they are not necessarily reflecting real interactions.

Classically, a network describes co-occurrence patterns between taxa. However, environmental variables can also be included to explore microbe-environment relationships. Additionally, considering the natural complexity of inter taxa relationships in an ecosystem, pairwise microbe-microbe correlations, derived from most current network analyses, need to be expanded to a higher order, such as three- or four-way correlations. The high-order microbial co-occurrence patterns may involve possible disruption or enhancement of another taxon to a pairwise relationship (Bairey et al. 2016). Such high-order co-occurrence patterns can also be unraveled by analyzing compositional data, as long as new and proper

Table 3 Examples of network analysis conducted in the marine ecosystem

Community	Ecosystem	Influential factors	Statistical methods	References
Network within a single system				
Bacteria and archaea	Sediment	–	Spearman's rank correlations	Liu et al. (2014b)
Bacteria	Sediment	–	Whittaker's index of associations	Buttigieg and Ramette (2015)
Bacteria	Seawater	–	WGCNA	Bryant et al. (2016)
Bacteria	Sediment	–	Spearman's rank correlations	Qiao et al. (2018)
Bacteria, archaea and microbial eukaryotes	Seawater	–	Spearman's rank correlations	Zhou et al. (2018)
Network across different systems				
Spatially resolved network				
Bacteria	Seawater	Water depth	Extended local similarity analysis	Cram et al. (2015)
Bacteria	Seawater	Particle size	SparCC	Milici et al. (2016)
Bacteria and archaea	Estuary	Water and sediment	Spearman's rank correlations	Wei et al. (2016)
Bacteria	Seawater	Water depth	Spearman's rank correlations	Cui et al. (2019)
Temporally resolved network				
Bacteria and archaea	Seawater (deep chlorophyll maximum)	Month	Local similarity analysis	Beman et al. (2011)
Bacteria	Seawater	Month	Extended local similarity analysis	Cram et al. (2015)
Bacteria, myoviruses and phytoplankton	Seawater	Daily to weekly	Local similarity analysis	Needham et al. (2017)
Protist	Seawater	Daily to weekly	Extended local similarity analysis	Berdjeb et al. (2018)
Bacteria and archaea	Seawater	Season	SparCC	Chafee et al. (2018)

statistical tools are developed (Bairey et al. 2016). Although co-occurrence patterns are not appropriate to imply accurate microbial interactions, their spatiotemporal dynamics hold the potential to affect the assembling processes and ecological roles of microbial communities.

Microbial biogeochemical roles and ecosystem functioning

Phylogeny and functional traits

To better understand microbial ecosystem functioning, creating links between an individual taxon and a specific function is required. However, it is infeasible to trace all the diverse biogeochemical processes and to relate them to taxa. This problem may be solved from the point of view of the microorganisms, as microbial functional capabilities have been found to connect strongly with phylogeny (Martiny et al. 2013; Zimmerman et al. 2013). For example, the presence of functional genes related to oxygenic photosynthesis, methane oxidation and sulfate reduction has been found to be highly phylogenetically conserved (Martiny et al. 2013). In coastal seawater, recent studies have shown that microorganisms that assimilated organic matter, including starch and glucose, were phylogenetically clustered (Bryson et al. 2017; Mayali and Weber 2018), reflecting phylogenetically

conserved resource partitioning in the coastal microbial loop (Bryson et al. 2017). Such phylogenetic conservation in substrate utilization supports similar distribution patterns (even at a broader taxonomic level; Philippot et al. 2010; Schmidt et al. 2016) and similar lifestyles among microbial relatives. Salazar et al. (2015) found that the particle-associated and free-living populations in the deep ocean had different phylogenetic origins. These observations enhance the possibility of inferring microbial functional traits with phylogenetic information.

However, growing evidence now suggests that variations in functional traits occur within closely related microbes, even at the loosely defined species level (Larkin and Martiny 2017). *Prochlorococcus*, the most abundant genus of photosynthetic organisms, diverges into high- and low-light-adapted ecotypes, which display different light-harvesting strategies (Bibby et al. 2003). Likewise, *Alteromonas macleodii*, a typical copiotrophic r-strategist, contains both surface and deep-sea ecotypes (Ivars-Martinez et al. 2008), which have been shown to differ substantially in their capacity to degrade algal polysaccharides (Neumann et al. 2015). On the other hand, microorganisms performing similar metabolic functions can be only distantly related (Martiny et al. 2015), the basic principle of functional redundancy. Louca et al. (2016, 2017) found high functional redundancy in both marine and plant-associated microbial communities,

implying that microbial functional traits are widely spread among microbial lineages.

Horizontal gene transfer and gene gain and loss likely result in the discrepancy in phylogeny-functional trait relationships (Fig. 2a). The type of functional trait is also an important contributor, with traits with slower evolutionary rates displaying clearer phylogenetically conserved patterns. Bryson et al. (2017) reported that traits of resource utilization in marine microbes, represented by patterns of substrate assimilation, were phylogenetically cohesive, whereas those of biosynthetic activity after assimilation were heterogeneous. Additionally, functional potential (gene presence) does not necessarily mean function execution (gene expression)

and functional capability is not equal to rate (the same functional trait can be different in rate). Although the capability of glucose incorporation is widespread in bacteria and displays a shallow phylogenetic clustering (Martiny et al. 2013), the glucose assimilation rates in soil bacteria have been observed to vary greatly across phylogeny and display a deep phylogenetic clustering (Morrissey et al. 2016). This finding highlights the importance of incorporating functional rates into phylogeny-function relatedness and raises the need to examine how quantitative changes in microbial function affect ecosystem functioning.
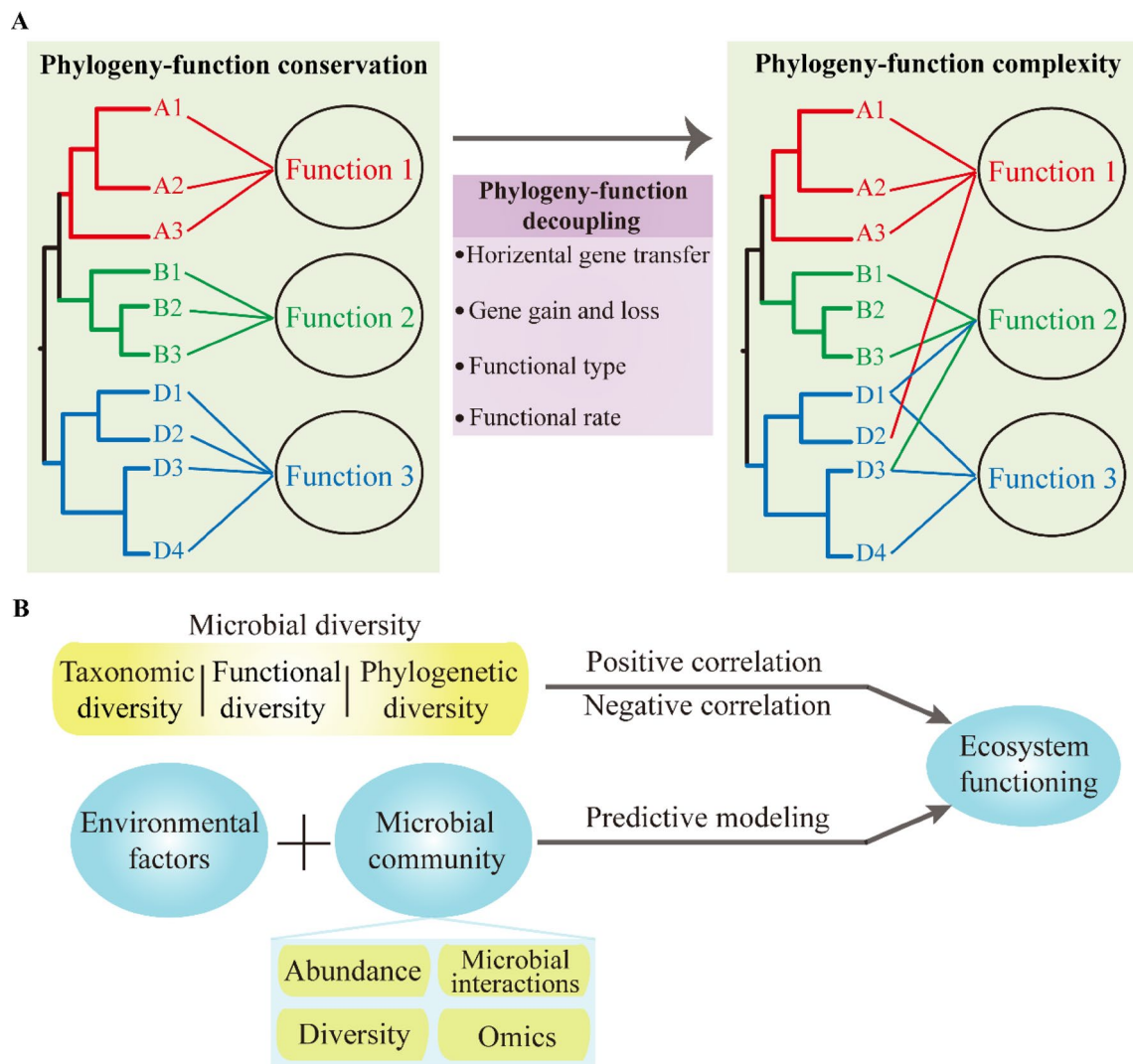
Fig. 2 Relationships among microbial phylogeny, diversity and ecosystem functioning: (a) the relationship between phylogeny and function is complex, although many functional traits are phylogenetically conserved. Possible reasons for the decoupling of phylogeny and function are presented; (b) microbial diversity (species richness, functional and phylogenetic diversity) is increasingly shown to be positively correlated to ecosystem functioning, whereas negative correlations are also observed. These correlations stimulate a great interest in predicting ecosystem functioning with microbial data. Inclusion of both environmental and microbial data (including microbial interactions and omic data) will enhance the power of ecosystem functioning prediction

Diversity and ecosystem functioning

There is growing evidence of a positive relationship between microbial diversity and ecosystem functioning (Cardinale et al. 2012; Delgado-Baquerizo et al. 2016b; Schnyder et al. 2018), although negative or no relationships have also been reported (Becker et al. 2012). Such relationships are derived primarily from studies on the terrestrial environment and have rarely been assessed for marine microbial communities. A study of microbial diversity-ecosystem functioning (DEF) relationship in marine surface water also supported a positive correlation, showing that a more phylogenetically diverse bacterial community had a greater level of ecosystem functioning (heterotrophic productivity measured by leucine incorporation; Galand et al. 2015). The enhancement of ecosystem functioning by increased biodiversity is thought to result from complementarity (minimal overlap) in resource use by functionally distinct taxa (Petchey and Gaston 2002) and/or through inter taxa facilitation (Hooper et al. 2005). Therefore, the relationship between diversity and ecosystem functioning is controlled by the niche-based mechanisms: differentiation in resource niche and selection effect (Krause et al. 2014).

The few studies that investigated the shape of the positive relationship between microbial diversity and ecosystem functioning have frequently uncovered a more linear relationship (Delgado-Baquerizo et al. 2016a) than the approaching-flat relationship seen for plants and animals (Cardinale et al. 2011). Such a linear relationship implies an indefinite increase of ecosystem functioning with increasing microbial diversity, challenging the idea of functional redundancy as mentioned above. In fact, Galand et al. (2018) provide evidence against the hypothesis of functional redundancy by showing a strong link between marine microbial community compositions and functional attributes using all the set of metagenomic reads. The authors emphasize the need to consider all functional aspects rather than relying only on known genes in investigating microbial DEF relationships. In addition, different processing rates seen in the same functional trait (Morrissey et al. 2016) may also provide opposing evidence against functional redundancy. However, these findings do not rule out the possibility for a partial functional redundancy, implicating that different types of functional traits may have different levels of redundancy. The idea of functional redundancy on the one hand can help to explain the high level of marine microbial diversity (different taxa are supported by a limited range of resources and conduct the same set of metabolic processes; Allison and Martiny 2008), while on the other hand can limit the extent of ecosystem functioning.

Diversity is composed of different components, including richness (taxonomic diversity), phylogeny (phylogenetic diversity) and function (functional diversity) (Fig. 2b).

Different types of diversity can inform distinct microbial DEF relationships. However, taxonomic diversity is the more frequently used proxy in inferring DEF relationships, compared to functional diversity and phylogenetic diversity. It has been reported that taxonomic diversity has relatively little impact on ecosystem functioning (Nielsen et al. 2011), while functional diversity was more correlated, mostly likely by determining ecological niches and inter taxa interactions (Hooper et al. 2005; Krause et al. 2014). Nevertheless, functional diversity is always difficult to measure (functional activity) and/or requires additional sequencing efforts to analyze (functional genes). Thus, phylogenetic diversity is increasingly implemented as a proxy of functional diversity, with the thought that many functional traits are phylogenetically conserved. Indeed, a positive correlation has been found between marine surface bacterial productivity and phylogenetic diversity of the active community; no similar association was found when taxonomic diversity (Shannon index) was analyzed (Galand et al. 2015). The findings of Galand et al. (2015) highlighted that ecosystem functioning is related to the active rather than the total community that contains dormant taxa. This provides an explanation for the more frequently observed negative and/or no relationships between phylogenetic diversity of the total community and ecosystem functioning (Goberna and Verdu 2018; Pérez-Valera et al. 2015; Severin et al. 2013). The relationship between phylogenetic diversity and ecosystem functioning also relates to taxon-specific functional capability and evolutionary history (Gravel et al. 2012). Under which conditions phylogenetic diversity can be used as a representative of functional diversity should be characterized further.

The microbial DEF relationship can be confounded by environmental variations, as environmental factors can exert influences on both diversity and ecosystem functioning. Orland et al. (2018) demonstrated that pH and organic matter quantity and quality explained as much variation in CO_2 production as did taxonomic diversity in lake sediments; these environmental factors exerted direct influences on ecosystem functioning due to their unrelatedness to taxonomic diversity. Comparatively, Delgado-Baquerizo et al. (2016b) found in a global set of soil samples that the DEF relationship was maintained when accounting for edaphic factors, which suggests that taxonomic diversity can exert influences on ecosystem functioning independently of environmental variabilities. A global survey of microbiome in seawater showed a decoupling of taxonomy and function, with the latter being more susceptible to environmental changes (Louca et al. 2016). Environmental conditions determine the availability of electronic donors/acceptors to microbes and shape the process of biogeochemical reactions. In addition to the environment, stochastic processes may also affect diversity and influence ecosystem functioning (Orland et al. 2018; Zhou et al. 2013). Overall, the positive DEF relationship

in microorganisms is supportive of phylogenetic conservation in functional traits. More effort is needed to discern the role of different diversity components and the role of deterministic and stochastic processes in determining ecosystem functioning.

Ecosystem functioning prediction

The abovementioned relationships have stimulated great interest in using microbial data to predict ecosystem functioning (statistical simulation instead of direct measurement) (Graham et al. 2016; Powell et al. 2015b) (Fig. 2b). Here, we focus on the interactions between community (diversity and abundance) and ecosystem functioning, although physiological properties can also be related (Wieder et al. 2013). Graham et al. (2016) synthesized 82 global datasets from different ecosystems to improve the predictive power of carbon and nitrogen processing rates by the inclusion of the microbial community data. They found that the addition of both compositional and diversity data could strengthen the predictive power, although this was not applicable to all datasets. Andersson et al. (2014) demonstrated, via structural equation models, that the model that included total bacterial abundance explained 54% of the variation in nitrogenase activity in coastal sediments and by replacing total bacterial abundance with cyanobacterial biomass it could increase the predictive power.

The explanatory power of microbial data in predictive models is always lower than that of abiotic factors (Graham et al. 2014, 2016; Powell et al. 2015b), consistent with the notion that environmental factors have direct impacts on ecosystem functioning. Moreover, under different environmental conditions (e.g., temperature, Dolan et al. 2017), the explanatory power of microbial data may change. However, this does not decrease the importance of microbial data in functional prediction. Recently, Zhang et al. (2018) found that in coastal sediments adding different copy number ratios of functional and rRNA genes into stepwise regression models substantially increased the predictive power of denitrification and anammox rates, although alpha diversity and gene abundance of involved bacteria were poorly correlated to the function potentials. In addition to abundance and diversity, it is also important to include microbial interactions to the predictive model in future studies (Fig. 2b). In summary, knowledge on the distribution patterns of microbial communities is indispensable for understanding their biogeochemical and ecological functions.

Inference of microbial activity

As stated above, a microbial community contains both active and dormant members. Dormant microbes constitute the seed bank of a community. Their repeated entrance to and exit from the seed bank decouple the active and total fractions (Jones and Lennon 2010; Fig. 3). Since RNA is a good indicator of whether or not a microbe is viable, rRNA based methods have been extensively used to characterize the active community, despite the difficulty in RNA extraction and the probable biases introduced by reverse transcription.

Mounting evidence suggests that there is a significant difference between the total (resident) and active fractions within a microbial community. Several abundant taxa are less active in the RNA pool, whereas some highly active taxa show low abundance or are almost absent in the DNA pool (Baldrian et al. 2012; Richa et al. 2017; Romanowicz et al. 2016; Sebastián et al. 2018). For example, *Cyanobacteria* and the SAR11 clade of *Alphaproteobacteria* are the most abundant microbes in the global surface ocean; whereas the former is always disproportionately active (16S rRNA:rDNA > 1), the latter tends to be less active (Campbell and Kirchman 2013; Hunt et al. 2013; Zhang et al. 2014). Further, a refined phylogenetic analysis showed that different ecotypes of the SAR11 clade varied in their 16S rRNA:rDNA ratios (Salter et al. 2015). A similar phenomenon was also shown for different ecotypes of MG-I (Hugoni et al. 2013). Within a microbial community, activity can also vary between rare and abundant taxa (Campbell et al. 2011; Richa et al. 2017). Richa et al. (2017) reported that more than 70% of the rare taxa in coastal seawater of the Mediterranean Sea had high 16S rRNA:rDNA ratios. To explain this decoupling of abundance and activity, Campbell et al. (2011) proposed that a substantial proportion of bacteria became active when their abundance decreased, indicating that high abundance may be a constraint factor for activity or that top-down processes i.e., grazing and virus lysis could stimulate activity. Seasonality (Hugoni et al. 2013), environmental factors, such as salinity (Campbell and Kirchman 2013), and lifestyle modes (free-living and particle-attached; Li et al. 2018) have also been reported as drivers for 16S rRNA:rDNA ratio variations.

The active and total microbial communities have been found to display contrasting biogeographic patterns and respond differently to environmental factors in seawater (Zhang et al. 2014). Environmental changes would generate uncomfortable conditions for active microbes, and the adaptation and successful establishment of active microbes in a new environment are difficult (Hanson et al. 2012). By contrast, the growth of several dormant microbes can be stimulated by changing environments, contributing to microbial community succession. Considering this, the active community is likely to display a stronger distance-decay relationship than the total community (Zhang et al. 2014). Nevertheless, our understanding of the assembling processes (relative role of deterministic and stochastic processes) of the total and active communities is limited. Zhang et al. (2014) also found that the active (dominated by negative correlations)
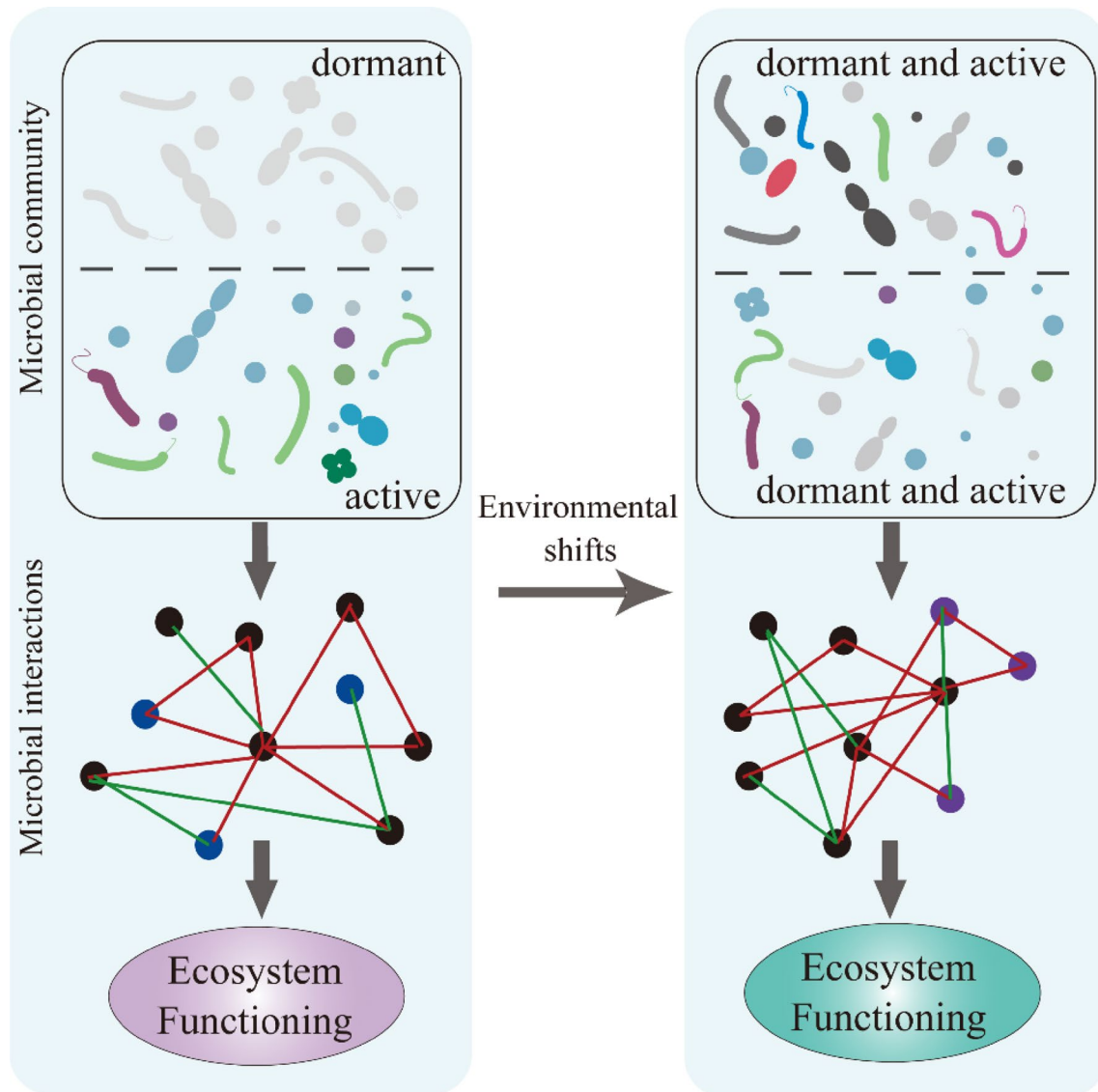
Fig. 3 Environmental shifts will change the relative abundance of the dormant and active fractions within a microbial community. Such changes will cause shifts in microbial co-occurrence patterns and finally affect ecosystem functioning

and total (dominated by positive correlations) bacterial communities exhibited different co-occurrence patterns. Frequent transitions of microbes between an active or dormant status under changing environments would lead to variations in co-occurrence patterns, causing significant alterations in ecosystem functioning (Fig. 3). As mentioned above, active microbes are directly linked to ecosystem functioning by actively carrying out biogeochemical reactions (Nannipieri et al. 2003). Thus, distinguishing the active fraction from the whole community would provide novel insights into patterns of microbial assembly, co-occurrence and DEF relationship. Moreover, the finding of a higher environmental sensitivity of active heterotrophs than active autotrophs in seawater

(Zhang et al. 2014) further raises the need to treat functionally different microbial groups separately.

Another possible utilization of the rRNA:rDNA ratio is to indicate potential growth rate (Campbell et al. 2011; Lankiewicz et al. 2016), as numerous microorganisms are yet to be cultivated and their growth rates are not able to be measured (Kirchman 2016). A high rRNA:rDNA ratio may imply a high growth rate. Thus, the lower proportion of SAR11 in the RNA than in the DNA pool, as mentioned above, may indicate a slow growth mode, although it may also be due to the low number of ribosomes per cell, given its small cell size. The slow growth rate of SAR11, however, is further verified by culture-based analyses (Lankiewicz et al. 2016). In comparison, copiotrophic taxa from *Alteromonas*

and the *Rosebacter* clade often display higher growth rates (Hunt et al. 2013; Lankiewicz et al. 2016). Noticeably, the rRNA:rDNA ratio has been reported to be not always effective in quantifying microbial growth rates, and protein synthesis potential has been proposed to be a more suitable interpretation (Blazewicz et al. 2013).

Insight into microbial diversification

The ribosomal gene sequences obtained from the marine environments using culture-independent methods have significantly expanded the microbial phylogenetic tree and public rRNA gene databases (the most popular are SILVA, Greengenes and RDP databases) (Brown et al. 2015; Delong 1992; Hug et al. 2016; Kubo et al. 2012). The microbial groups that have no close isolates at the time of discovering their rRNA gene sequences are often indicated by informal nomenclature, such as SAR11, MG-I and Miscellaneous Crenarchaeota Group (MCG). Further, in-depth phylogenetic analyses of the rRNA gene sequences have clustered many microbial groups, such as MG-I and MCG, into subclades (Kubo et al. 2012; Liu et al. 2014a). These subclades have been shown to respond differently to environmental changes and exhibit different habitat preferences (Lazar et al. 2016; Liu et al. 2014a). Similar niche differentiations are also shown for marine bacteria sharing high 16S rRNA gene similarities (e.g., *Alteromonas macleodii* as mentioned above). These observations suggest that environmental selection plays a crucial role in the process of microbial evolution. In addition to rRNA genes, functional genes can be used to investigate the diversification of functional groups. For example, Alves et al. (2018) recently elucidated the global frequency, phylogenetic diversity and habitat specificity of ammonia-oxidizing archaea using *amoA*, the gene encoding the active-site subunit of ammonia monooxygenase. These results highlight the importance of using community data derived from molecular marker genes to investigate the phylogenetic relationships and to infer the evolutionary histories of different microbial taxa.

However, a single marker gene is not always effective in differentiating phylogenetically close relatives. In this context, the multilocus sequence typing (MLST) analysis, which takes the phylogenetic information of several (usually 5–7) housekeeping genes, has been shown to provide a better resolution (Enright and Spratt 1999). In recent years, the increasing number of public metagenomic/genomic sequences has provided an excellent resource for determining phylogenetic relatedness among microbes with a high confidence level (Brown et al. 2015; Hug et al. 2016). These phylogenomic analyses have enabled the designation of many novel microbial phyla, including the well-known *Thaumarchaeota* (containing MG-I) and *Bathyarchaeota* (MCG) (Adam et al.

2017). Spang et al. (2015) proposed a new archaeal phylum *Lokiarchaeota* through phylogenomic inference. This phylum contains many eukaryotic signatures, significantly contributing to our understanding of the evolution of life.

Future perspective

The prevalence of studies on microbial communities in the marine environment benefits from the development of new sequencing technologies. The resulting increase in sequencing depth facilities the obtaining of accurate insights into microbial community structure, and in particular enhances the capability of distinguishing rare subcommunities from sequencing errors. In the era of big data, the development of sequencing technologies calls for a simultaneous introduction of new statistical methods and tools, which would help to discern microbial assembling processes and unravel high-order (three- or four-way) microbe-microbe interactions that most likely occur in natural environments. Although microbe-microbe interactions, to date derived mainly from co-occurrence networks, are important in structuring microbial communities and mediating ecological functions, such linkages are still largely elusive. Suitable analytic tools are required to construct a model that bridges the gap between microbial interactions and ecosystem functioning.

Owing to the difficulty in measuring microbial functions, inclusion of community data in the predictive modeling of ecosystem functioning is receiving increasing attention. This raises associated questions such as whether different functional types exhibit similar sensitivity to community data and vice versa, and whether relationships between community data and ecosystem functioning vary spatially and temporally. In addition, the decoupling of phylogeny and function may act as an obstacle for such modeling, which calls for a resolved phylogenetic clustering and taxon-specific functional characterization. Omics, such as metagenome, single-cell genome and metatranscriptome, in parallel with advanced analytic tools, provide great opportunities to link a community, and further to link an individual to function (Fig. 2b).

In this review, we highlight the significance of separating a whole microbial community into subgroups according to different standards. The abundance-dependent grouping has led to the realization that rare subcommunities, the members of which often show high activity, have the potential to increase the power of ecosystem functioning prediction. By contrast, functional-, and other traits-dependent grouping have received less attention. After it has been established what exists, it is then necessary to determine what is alive and whether the active taxa follow patterns seen for the total community, and how active taxa contribute to ecosystem functioning. We also need to know if functional traits of the

active community are phylogenetically organized, because such information is important for constructing a direct linkage between community and ecosystem functioning. In summary, subdivision of the whole community will provide accurate and novel insights into relationships between microbial diversity, interaction and ecosystem functioning.

Author's Contribution XZ conceived, provided the idea to this work and edited the manuscript; JL, ZM and XL wrote the manuscript; JL prepared the figures and tables. All authors approved the final manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Animal and human rights statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

Adam PS, Borrel G, Brochier-Armanet C, Gribaldo S (2017) The growing tree of archaea: new perspectives on their diversity, evolution and ecology. ISME J 11:2407–2425

Allison SD, Martiny JBH (2008) Resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci USA 105:11512–11519

Alves RJE, Minh BQ, Urich T, von Haeseler A, Schleper C (2018) Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA genes. Nat Commun 9:1517

Andersson B, Sundbäck K, Hellman M, Hallin S, Alsterberg C (2014) Nitrogen fixation in shallow-water sediments: spatial distribution and controlling factors. Limnol Oceanogr 59:1932–1944

Aßhauer KP, Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. Bioinformatics 31:2882–2884

Baas-Becking LGM (1934) Geobiologie of inleiding tot de milieukunde. W.P. van Stockum & Zoon, Den Haag

Bairey E, Kelsic ED, Kishony R (2016) High-order species interactions shape ecosystem diversity. Nat Commun 7:12285

Baldrian P, Kolarik M, Stursova M, Kopecky J, Valaskova V, Vetrovsky T, Zifcakova L, Snajdr J, Ridl J, Vlcek C, Voriskova J (2012) Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. ISME J 6:248–258

Banerjee S, Schlaeppi K, van der Heijden MGA (2018) Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol 16:567–576

Barberán A, Bates ST, Casamayor EO, Fierer N (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J 6:343–351

Becker J, Eisenhauer N, Scheu S, Jousset A (2012) Increasing antagonistic interactions cause bacterial communities to collapse at high diversity. Ecol Lett 15:468–474

Beman JM, Steele JA, Fuhrman JA (2011) Co-occurrence patterns for abundant marine archaeal and bacterial lineages in the deep chlorophyll maximum of coastal California. ISME J 5:1077–1085

Berdjeb L, Parada A, Needham DM, Fuhrman JA (2018) Short-term dynamics and interactions of marine protist communities during the spring-summer transition. ISME J 12:1907–1917

Berry D, Widder S (2014) Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Front Microbiol 5:219

Bibby TS, Mary I, Nield J, Partensky F, Barber J (2003) Low-light-adapted *Prochlorococcus* species possess specific antennae for each photosystem. Nature 424:1051–1054

Blazewicz SJ, Barnard RL, Daly RA, Firestone MK (2013) Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. ISME J 7:2061–2068

Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF (2015) Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523:208–211

Bryant JA, Aylward FO, Eppley JM, Karl DM, Church MJ, DeLong EF (2016) Wind and sunlight shape microbial diversity in surface waters of the North Pacific Subtropical Gyre. ISME J 10:1308–1322

Bryson S, Li Z, Chavez F, Weber PK, Pett-Ridge J, Hettich RL, Pan C, Mayali X, Mueller RS (2017) Phylogenetically conserved resource partitioning in the coastal microbial loop. ISME J 11:2781–2792

Buttigieg PL, Ramette A (2015) Biogeographic patterns of bacterial microdiversity in Arctic deep-sea sediments (HAUSGARTEN, Fram Strait). Front Microbiol 5:660

Campbell BJ, Kirchman DL (2013) Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. ISME J 7:210–220

Campbell BJ, Yu LY, Heidelberg JF, Kirchman DL (2011) Activity of abundant and rare bacteria in a coastal ocean. Proc Natl Acad Sci USA 108:12776–12781

Cardinale BJ, Matulich KL, Hooper DU, Byrnes JE, Duffy E, Gamfeldt L, Balvanera P, O'Connor MI, Gonzalez A (2011) The functional role of producer diversity in ecosystems. Am J Bot 98:572–592

Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Narwani A, Mace GM, Tilman D, Wardle DA, Kinzig AP, Daily GC, Loreau M, Grace JB, Larigauderie A, Srivastava DS, Naeem S (2012) Biodiversity loss and its impact on humanity. Nature 486:59–67

Caruso T, Chan YK, Lacap DC, Lau MCY, Mckay CP, Pointing SB (2011) Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. ISME J 5:1406–1413

Chafee M, Fernandez-Guerra A, Buttigieg PL, Gerdts G, Eren AM, Teeling H, Amann RI (2018) Recurrent patterns of microdiversity in a temperate coastal marine environment. ISME J 12:237–252

Chase JM, Leibold MA (2003) Ecological niches: linking classical and contemporary approaches. University of Chicago Press, Chicago

Chase JM, Myers JA (2011) Disentangling the importance of ecological niches from stochastic processes across scales. Philos Trans R Soc B 366:2351–2363

Chave J (2004) Neutral theory and community ecology. Ecol Lett 7:241–253

Chen W, Pan Y, Yu L, Yang J, Zhang W (2017) Patterns and processes in marine microeukaryotic community biogeography from Xiamen coastal waters and intertidal sediments, southeast China. Front Microbiol 8:1912

Chow CET, Sachdeva R, Cram JA, Steele JA, Needham DM, Patel A, Parada AE, Fuhrman JA (2013) Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. ISME J 7:2259–2273

Cram JA, Xia LC, Needham DM, Sachdeva R, Sun F, Fuhrman JA (2015) Cross-depth analysis of marine bacterial networks suggests downward propagation of temporal changes. ISME J 9:2573–2586

Cui Y, Chun SJ, Baek SH, Lee M, Kim Y, Lee HG, Ko SR, Hwang S, Ahn CY, Oh HM (2019) The water depth-dependent co-occurrence patterns of marine bacteria in shallow and dynamic Southern Coast, Korea. Sci Rep 9:9176

Dai W, Zhang J, Tu Q, Deng Y, Qiu Q, Xiong J (2017) Bacterioplankton assembly and interspecies interaction indicating increasing coastal eutrophication. Chemosphere 177:317–325

Delgado-Baquerizo M, Giaramida L, Reich PB, Khachane AN, Hamonts K, Edwards C, Lawton LA, Singh BK (2016a) Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. J Ecol 104:936–946

Delgado-Baquerizo M, Maestre FT, Reich PB, Jeffries TC, Gaitan JJ, Encinar D, Berdugo M, Campbell CD, Singh BK (2016b) Microbial diversity drives multifunctionality in terrestrial ecosystems. Nat Commun 7:10541

DeLong EF (1992) Archaea in coastal marine environments. Proc Natl Acad Sci USA 89:5685–5689

Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF (2015) Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. Proc Natl Acad Sci USA 112:E1326–E1332

Dolan KL, Pena J, Allison SD, Martiny JBH (2017) Phylogenetic conservation of substrate use specialization in leaf litter bacteria. PLoS ONE 12:e0174472

Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH (2010) Relative roles of niche and neutral processes in structuring a soil microbial community. ISME J 4:337–345

Enright MC, Spratt BG (1999) Multilocus sequence typing. Trends Microbiol 7:482–487

Evans S, Martiny JBH, Allison SD (2017) Effects of dispersal and selection on stochastic assembly in microbial communities. ISME J 11:176–185

Galand PE, Casamayor EO, Kirchman DL, Lovejoy C (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. Proc Natl Acad Sci USA 106:22427–22432

Galand PE, Salter I, Kalenitchenko D (2015) Ecosystem productivity is associated with bacterial phylogenetic distance in surface marine waters. Mol Ecol 24:5785–5795

Galand PE, Pereira O, Hochart C, Auguet JC, Debroas D (2018) A strong link between marine microbial community composition and function challenges the idea of functional redundancy. ISME J 12:2470–2478

Goberna M, Verdu M (2018) Phylogenetic-scale disparities in the soil microbial diversity-ecosystem functioning relationship. ISME J 12:2152–2162

Gong J, Shi F, Ma B, Dong J, Pachiadaki M, Zhang XL, Edgcomb VP (2015) Depth shapes alpha- and beta-diversities of microbial eukaryotes in surficial sediments of coastal ecosystems. Environ Microbiol 17:3722–3737

Graham EB, Wieder WR, Leff JW, Weintraub SR, Townsend AR, Cleveland CC, Philippot L, Nemergut DR (2014) Do we need to understand microbial communities to predict ecosystem function? A comparison of statistical models of nitrogen cycling processes. Soil Biol Biochem 68:279–282

Graham EB, Knelman JE, Schindlbacher A, Siciliano S, Breulmann M, Yannarell A, Bemans JM, Abell G, Philippot L, Prosser J, Foulquier A, Yuste JC, Glanville HC, Jones DL, Angel F, Salminen J, Newton RJ, Burgmann H, Ingram LJ, Hamer U et al (2016) Microbes as engines of ecosystem function: when does community structure enhance predictions of ecosystem processes? Front Microbiol 7:214

Graham EB, Crump AR, Resch CT, Fansler S, Arntzen E, Kennedy DW, Fredrickson JK, Stegen JC (2017) Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. Environ Microbiol 19:1552–1567

Gravel D, Bell T, Barbera C, Combe M, Pommier T, Mouquet N (2012) Phylogenetic constraints on ecosystem functioning. Nat Commun 3:1117

Gutknecht JLM, Field CB, Balser TC (2012) Microbial communities and their responses to simulated global change fluctuate greatly over multiple years. Glob Change Biol 18:2256–2269

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

Hooper DU, Chapin FS, Ewel JJ, Hector A, Inchausti P, Lavorel S, Lawton JH, Lodge DM, Loreau M, Naeem S, Schmid B, Setala H, Symstad AJ, Vandermeer J, Wardle DA (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecol Monogr 75:3–35

Horner-Devine MC, Lage M, Hughes JB, Bohannan BJ (2004) A taxaarea relationship for bacteria. Nature 432:750–753

Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography (MPB-32), vol 32. Princeton University Press, Princeton

Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K, Suzuki Y, Dudek N, Relman DA, Finstad KM, Amundson R, Thomas BC, Banfield JF (2016) A new view of the tree of life. Nat Microbiol 1:16048

Hugoni M, Taib N, Debroas D, Domaizon I, Dufournel IJ, Bronner G, Salter I, Agogué H, Mary I, Galand PE (2013) Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. Proc Natl Acad Sci USA 110:6004–6009

Hunt DE, Lin Y, Church MJ, Karl DM, Tringe SG, Izzo LK, Johnson ZI (2013) Relationship between abundance and specific activity of bacterioplankton in open ocean surface waters. Appl Environ Microbiol 79:177–184

Hutchins DA, Fu F (2017) Microorganisms and ocean global change. Nat Microbiol 2:17058

Inagaki F, Suzuki M, Takai K, Oida H, Sakamoto T, Aoki K, Nealson KH, Horikoshi K (2003) Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. Appl Environ Microbiol 69:7224–7235

Ivars-Martinez E, D'Auria G, Rodriguez-Valera F, Sanchez-Porro C, Ventosa A, Joint I, Muhling M (2008) Biogeography of the ubiquitous marine bacterium *Alteromonas macleodii* determined by multilocus sequence analysis. Mol Ecol 17:4092–4106

Jeraldo P, Sipos M, Chia N, Brulc JM, Dhillon AS, Konkel ME, Larson CL, Nelson KE, Qu A, Schook LB, Yang F, White BA, Goldenfeld N (2012) Quantification of the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes. Proc Natl Acad Sci USA 109:9692–9698

Jia X, Dini-Andreote F, Salles JF (2018) Community assembly processes of the microbial rare biosphere. Trends Microbiol 26:738–747

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

Ju F, Xia Y, Guo F, Wang Z, Zhang T (2014) Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. Environ Microbiol 16:2421–2432

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

Krause S, Le Roux X, Niklaus PA, Van Bodegom PM, Lennon JT, Bertilsson S, Grossart HP, Philippot L, Bodelier PLE (2014) Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. Front Microbiol 5:251

Kubo K, Lloyd KG, Biddle JF, Amann R, Teske A, Knittel K (2012) Archaea of the miscellaneous crenarchaeotal group are abundant, diverse and widespread in marine sediments. ISME J 6:1949–1965

Langenheder S, Berga M, Östman Ö, Székely AJ (2012) Temporal variation of β-diversity and assembly mechanisms in a bacterial metacommunity. ISME J 6:1107–1114

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Thurber RLV, Knight R, Beiko RG, Huttenhower C (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 31:814–821

Lankiewicz TS, Cottrell MT, Kirchman DL (2016) Growth rates and rRNA content of four marine bacteria in pure cultures and in the Delaware estuary. ISME J 10:823–832

Larkin AA, Martiny AC (2017) Microdiversity shapes the traits, niche space, and biogeography of microbial taxa. Environ Microbiol Rep 9:55–70

Layeghifard M, Hwang DM, Guttman DS (2017) Disentangling interactions in the microbiome: a network perspective. Trends Microbiol 25:217–228

Lazar CS, Baker BJ, Seitz K, Hyde AS, Dick GJ, Hinrichs KU, Teske AP (2016) Genomic evidence for distinct carbon substrate preferences and ecological niches of *Bathyarchaeota* in estuarine sediments. Environ Microbiol 18:1200–1211

Leibold MA (1995) The niche concept revisited-mechanistic models and community context. Ecology 76:1371–1382

Leibold MA, McPeek MA (2006) Coexistence of the niche and neutral perspectives in community ecology. Ecology 87:1399–1410

Li J, Salam N, Wang P, Chen L, Jiao J, Li X, Xian W, Han M, Fang B, Mou X, Li W (2018) Discordance between resident and active bacterioplankton in free-living and particle-associated communities in estuary ecosystem. Microb Ecol 76:637–647

Liang Y, Zhao H, Deng Y, Zhou J, Li G, Sun B (2016) Long-term oil contamination alters the molecular ecological networks of soil microbial functional genes. Front Microbiol 7:60

Liao J, Cao X, Zhao L, Wang J, Gao Z, Wang MC, Huang Y (2016) The importance of neutral and niche processes for bacterial community assembly differs between habitat generalists and specialists. FEMS Microbiol Ecol 92:fiw174

Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, Chaffron S, Ignacio-Espinosa JC, Roux S, Vincent F, Bittner L, Darzi Y, Wang J, Audic S, Berline L, Bontempi G, Cabello AM, Coppola L, Cornejo-Castillo FM, d'Ovidio F et al (2015) Ocean plankton. Determinants of community structure in the global plankton interactome. Science 348:1262073

Liu J, Yu S, Zhao M, He B, Zhang X-H (2014a) Shifts in archaea-plankton community structure along ecological gradients of Pearl Estuary. FEMS Microbiol Ecol 90:424–435

Liu J, Yang H, Zhao M, Zhang X-H (2014b) Spatial distribution patterns of benthic microbial communities along the Pearl Estuary, China. Syst Appl Microbiol 37:578–589

Liu J, Fu B, Yang H, Zhao M, He B, Zhang X-H (2015a) Phylogenetic shifts of bacterioplankton community composition along the Pearl Estuary: the potential impact of hypoxia and nutrients. Front Microbiol 6:64

Liu L, Yang J, Yu Z, Wilkinson DM (2015b) The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. ISME J 9:2068–2077

Liu J, Zheng Y, Lin H, Wang X, Li M, Liu Y, Yu M, Zhao M, Pedentchouk N, Lea-Smith DJ, Todd JD, Magill CR, Zhang WJ, Zhou S, Song D, Zhong H, Xin Y, Yu M, Tian J, Zhang X-H (2019) Proliferation of hydrocarbon-degrading microbes at the bottom of the Mariana Trench. Microbiome 7:47

Logares R, Tesson SV, Canback B, Pontarp M, Hedlund K, Rengefors K (2018) Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. Environ Microbiol 20:2231–2240

Louca S, Parfrey LW, Doebeli M (2016) Decoupling function and taxonomy in the global ocean microbiome. Science 353:1272–1277

Louca S, Jacques SMS, Pires APF, Leal JS, Srivastava DS, Parfrey LW, Farjalla VF, Doebeli M (2017) High taxonomic variability despite stable functional structure across microbial communities. Nat Ecol Evol 1:15

Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. Proc Natl Acad Sci USA 104:11436–11440

Ma B, Wang H, Dsouza M, Lou J, He Y, Dai Z, Brookes PC, Xu J, Gilbert JA (2016) Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. ISME J 10:1891–1901

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Ovreas L, Reysenbach AL, Smith VH, Staley JT (2006) Microbial biogeography: putting microorganisms on the map. Nat Rev Microbiol 4:102–112

Martiny AC, Treseder K, Pusch G (2013) Phylogenetic conservatism of functional traits in microorganisms. ISME J 7:830–838

Martiny JB, Jones SE, Lennon JT, Martiny AC (2015) Microbiomes in light of traits: a phylogenetic perspective. Science 350:aac9323

Mayali X, Weber PK (2018) Quantitative isotope incorporation reveals substrate partitioning in a coastal microbial community. FEMS Microbiol Ecol 94:fiy047

Milici M, Deng ZL, Tomasch J, Decelle J, Wos-Oxley ML, Wang H, Jauregui R, Plumeier I, Giebel HA, Badewien TH, Wurst M, Pieper DH, Simon M, Wagner-Dobler I (2016) Co-occurrence analysis of microbial taxa in the Atlantic Ocean reveals high connectivity in the free-living bacterioplankton. Front Microbiol 7:649

Mo Y, Zhang W, Yang J, Lin Y, Yu Z, Lin S (2018) Biogeographic patterns of abundant and rare bacterioplankton in three subtropical bays resulting from selective and neutral processes. ISME J 12:2198–2210

Morrissey EM, Mau RL, Schwartz E, Caporaso JG, Dijkstra P, van Gestel N, Koch BJ, Liu CM, Hayer M, McHugh TA, Marks JC, Price LB, Hungate BA (2016) Phylogenetic organization of bacterial activity. ISME J 10:2336–2340

Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G (2003) Microbial diversity and soil functions. Eur J Soil Sci 54:655–670

Needham DM, Sachdeva R, Fuhrman JA (2017) Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. ISME J 11:1614–1629

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, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S (2013) Patterns and processes of microbial community assembly. Microbiol Mol Biol Rev 77:342–356

Neumann AM, Balmonte JP, Berger M, Giebel H-A, Arnosti C, Voget S, Simon M, Brinkhoff T, Wietz M (2015) Different utilization of alginate and other algal polysaccharides by marine *Alteromonas macleodii* ecotypes. Environ Microbiol 17:3858–3868

Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45:167–256

Nielsen UN, Ayres E, Wall DH, Bardgett RD (2011) Soil biodiversity and carbon cycling: a review and synthesis of studies examining diversity-function relationships. Eur J Soil Sci 62:105–116

Orland C, Emilson EJS, Basiliko N, Mykytczuk NCS, Gunn JM, Tanentzap AJ (2018) Microbiome functioning depends on individual and interactive effects of the environment and community structure. ISME J 13:1–11

Pedrós-Alió C (2006) Marine microbial diversity: can it be determined? Trends Microbiol 14:257–263

Pedrós-Alió C (2012) The rare bacterial biosphere. Annu Rev Mar Sci 4:449–466

Pérez-Valera E, Goberna M, Verdu M (2015) Phylogenetic structure of soil bacterial communities predicts ecosystem functioning. FEMS Microbiol Ecol 91:fiv031

Petchey OL, Gaston KJ (2002) Extinction and the loss of functional diversity. Proc R Soc Lond B 269:1721–1727

Philippot L, Andersson SGE, Battin TJ, Prosser JI, Schimel JP, Whitman WB, Hallin S (2010) The ecological coherence of high bacterial taxonomic ranks. Nat Rev Microbiol 8:523–529

Powell JR, Karunaratne S, Campbell CD, Yao H, Robinson L, Singh BK (2015a) Deterministic processes vary during community assembly for ecologically dissimilar taxa. Nat Commun 6:8444

Powell JR, Welsh A, Hallin S (2015b) Microbial functional diversity enhances predictive models linking environmental parameters to ecosystem properties. Ecology 96:1985–1993

Preston FW (1948) The commonness, and rarity, of species. Ecology 29:254–283

Qiao Y, Liu J, Zhao M, Zhang X-H (2018) Sediment depth-dependent spatial variations of bacterial communities in mud deposits of the eastern China marginal seas. Front Microbiol 9:1128

Reji L, Tolar BB, Smith JM, Chavez FP, Francis CA (2019) Differential co-occurrence relationships shaping ecotype diversification within *Thaumarchaeota* populations in the coastal ocean water column. ISME J 13:1144–1158

Richa K, Balestra C, Piredda R, Benes V, Borra M, Passarelli A, Margiotta F, Saggiomo M, Biffali E, Sanges R, Scanlan DJ, Casotti R (2017) Distribution, community composition, and potential metabolic activity of bacterioplankton in an urbanized mediterranean sea coastal zone. Appl Environ Microbiol 83:e00494–17

Roguet A, Laigle GS, Therial C, Bressy A, Soulignac F, Catherine A, Lacroix G, Jardillier L, Bonhomme C, Lerch TZ (2015) Neutral community model explains the bacterial community assembly in freshwater lakes. FEMS Microbiol Ecol 91:fiv125

Romanowicz KJ, Freedman ZB, Upchurch RA, Argiroff WA, Zak DR (2016) Active microorganisms in forest soils differ from the total community yet are shaped by the same environmental factors: the influence of pH and soil moisture. FEMS Microbiol Ecol 92:fiw149

Salazar G, Cornejo-Castillo FM, Borrull E, Diez-Vives C, Lara E, Vaque D, Arrieta JM, Duarte CM, Gasol JM, Acinas SG (2015) Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. Mol Ecol 24:5692–5706

Salter I, Galand PE, Fagervold SK, Lebaron P, Obernosterer I, Oliver MJ, Suzuki MT, Tricoire C (2015) Seasonal dynamics of active SAR11 ecotypes in the oligotrophic Northwest Mediterranean Sea. ISME J 9:347–360

Schmidt ML, White JD, Denef VJ (2016) Phylogenetic conservation of freshwater lake habitat preference varies between abundant bacterioplankton phyla. Environ Microbiol 18:1212–1226

Schnyder E, Bodelier PLE, Hartmann M, Henneberger R, Niklaus PA (2018) Positive diversity-functioning relationships in model communities of methanotrophic bacteria. Ecology 99:714–723

Sebastián M, Auguet JC, Restrepo-Ortiz CX, Sala MM, Marrase C, Gasol JM (2018) Deep ocean prokaryotic communities are remarkably malleable when facing long-term starvation. Environ Microbiol 20:713–723

Severin I, Ostman O, Lindstrom ES (2013) Variable effects of dispersal on productivity of bacterial communities due to changes in functional trait composition. PLoS ONE 8:e80825

Singh BK, Bardgett RD, Smith P, Reay DS (2010) Microorganisms and climate change: terrestrial feedbacks and mitigation options. Nat Rev Microbiol 8:779–790

Spang A, Saw JH, Jorgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJG (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature 521:173–179

Stegen JC, Lin XJ, Fredrickson JK, Chen XY, Kennedy DW, Murray CJ, Rockhold ML, Konopka A (2013) Quantifying community assembly processes and identifying features that impose them. ISME J 7:2069–2079

Stegen JC, Lin XJ, Fredrickson JK, Konopka AE (2015) Estimating and mapping ecological processes influencing microbial community assembly. Front Microbiol 6:370

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, Djahanschiri B, Zeller G, Mende DR, Alberti A, Cornejo-Castillo FM, Costea PI, Cruaud C, d'Ovidio F, Engelen S, Ferrera I, Gasol JM, Guidi L, Hildebrand F, Kokoszka F et al (2015) Ocean plankton. Structure and function of the global ocean microbiome. Science 348:1261359

Tripp HJ, Schwalbach MS, Meyer MM, Kitner JB, Breaker RR, Giovannoni SJ (2009) Unique glycine-activated riboswitch linked to glycine-serine auxotrophy in SAR11. Environ Microbiol 11:230–238

Volkov I, Banavar JR, Hubbell SP, Maritan A (2003) Neutral theory and relative species abundance in ecology. Nature 424:1035–1037

Wang J, Shen J, Wu Y, Tu C, Soininen J, Stegen JC, He J, Liu X, Zhang L, Zhang E (2013) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. ISME J 7:1310–1321

Wang K, Hu H, Yan H, Hou D, Wang Y, Dong P, Zhang D (2019) Archaeal biogeography and interactions with microbial community across complex subtropical coastal waters. Mol Ecol. https://doi.org/10.1111/mec.15105

Wei G, Li M, Li F, Li H, Gao Z (2016) Distinct distribution patterns of prokaryotes between sediment and water in the Yellow River estuary. Appl Microbiol Biotechnol 100:9683–9697

Weiss S, Van Treuren W, Lozupone C, Faust K, Friedman J, Deng Y, Xia LC, Xu ZZ, Ursell L, Alm EJ, Birmingham A, Cram JA, Fuhrman JA, Raes J, Sun FZ, Zhou JZ, Knight R (2016) Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. ISME J 10:1669–1681

Wieder WR, Bonan GB, Allison SD (2013) Global soil carbon projections are improved by modelling microbial processes. Nat Clim Change 3:909–912

Wu W, Logares R, Huang B, Hsieh CH (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

Wu W, Lu H, Sastri A, Yeh YC, Gong G, Chou W, Hsieh CH (2018) Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. ISME J 12:485–494

Zhang Y, Zhao Z, Dai M, Jiao N, Herndl GJ (2014) Drivers shaping the diversity and biogeography of total and active bacterial communities in the South China Sea. Mol Ecol 23:2260–2274

Zhang X, Zhang Q, Yang A, Hou L, Zheng Y, Zhai W, Gong J (2018) Incorporation of microbial functional traits in biogeochemistry models provides better estimations of benthic denitrification and anammox rates in coastal oceans. J Geophys Res Biogeosci 123:3331–3352

Zhou J, Ning D (2017) Stochastic community assembly: Does it matter in microbial ecology? Microbiol Mol Biol Rev 81:e00002–e00017

Zhou J, Liu W, Deng Y, Jiang Y, Xue K, He Z, Van Nostrand JD, Wu L, Yang Y, Wang A (2013) Stochastic assembly leads to alternative communities with distinct functions in a bioreactor microbial community. Mbio 4:e00584–12

Zhou J, Song X, Zhang CY, Chen GF, Lao YM, Jin H, Cai ZH (2018) Distribution patterns of microbial community structure along a 7000-mile latitudinal transect from the mediterranean sea across the Atlantic Ocean to the Brazilian coastal sea. Microb Ecol 76:592–609

Zimmerman AE, Martiny AC, Allison SD (2013) Microdiversity of extracellular enzyme genes among sequenced prokaryotic genomes. ISME J 7:1187–1199