Ecological Processes

RESEARCH

Open Access

Biogeographic shifts in the microbial co-occurrence network features of three domains across complex environmental gradients in subtropical coastal waters

Dandi Hou[1,2], Huizhen Yan[1], Huaying Lin[1], Huajun Zhang[1,2], Demin Zhang[1,2] and Kai Wang[1,2*]

Abstract

Background Bacteria, Archaea, and Microeukaryotes comprise taxonomic domains that interact in mediating biogeochemical cycles in coastal waters. Many studies have revealed contrasting biogeographic patterns of community structure and assembly mechanisms in microbial communities from different domains in coastal ecosystems; however, knowledge of specific biogeographic patterns on microbial co-occurrence relationships across complex coastal environmental gradients remains limited. Using a dense sampling scheme at the regional scale, SSU rRNA gene amplicon sequencing, and network analysis, we investigated intra- and inter-domain co-occurrence relationships and network topology-based biogeographic patterns from three microbial domains in coastal waters that show environmental gradients across the inshore-nearshore-offshore continuum in the East China Sea.

Results Overall, we found the highest complexity and connectivity in the bacterial network, the highest modularity in the archaeal network, and the lowest complexity, connectivity, and modularity in the microeukaryotic network. Although microbial co-occurrence networks from the three domains showed distinct topological features, they exhibited a consistent biogeographic pattern across the inshore-nearshore-offshore continuum. Specifically, the nearshore zones with intermediate levels of terrestrial impacts reflected by multiple environmental factors (including water temperature, salinity, pH, dissolved oxygen, and nutrient-related parameters) had a higher intensity of microbial co-occurrence for all three domains. In contrast, the intensity of microbial co-occurrence was weaker in both the inshore and the offshore zones at the two ends of the environmental gradients. Archaea occupied a central position in the microbial inter-domain co-occurrence network. In particular, members of the Thaumarchaeota Marine Group I (MGI, now placed within the Family Nitrosopumilaceae of the Phylum Thermoproteota) appeared to be the hubs in the biogeographic shift between inter-domain network modules across environmental gradients.

Conclusions Our work offers new insights into microbial biogeography by integrating network features into biogeographic patterns, towards a better understanding of the potential of microbial interactions in shaping biogeographic patterns of coastal marine microbiota.

Keywords Co-occurrence, Biogeography, Bacteria, Archaea, Microeukaryote, Environmental gradient

*Correspondence:
Kai Wang
wangkai@nbu.edu.cn
Full list of author information is available at the end of the article

Background

The surface marine microbiota is a pivotal foundation of global biogeochemical cycles (Falkowski 2012; Logares et al. 2020). Marine planktonic domains of Bacteria, Archaea, and eukaryotic microorganisms (hereafter Microeukaryotes) have diverse interspecific interactions, including mutualism, competition, or predation, (or none) (Abreu et al. 2022; Faust and Raes 2012), which structure marine microbial food webs (Lima-Mendez et al. 2015). Uncovering microbial association patterns within and among these three domains can provide insights into mechanisms governing community assembly and maintaining biodiversity in the complex microbiome (Fuhrman et al. 2015).

Over recent decades, a surge in marine microbial ecology research has provided a wealth of understanding of the patterns and determinants of microbial community dynamics over time and space (Deutschmann et al. 2024). These highlight the importance of characterizing biogeographic patterns of microbial diversity to understand microbial contributions and regulation of ecosystem function at or beyond the regional scale. This is particularly true for coastal microbiota as they play critical roles in regulating biogeochemical cycling at the land-sea interface (Trevathan-Tackett et al. 2019). Primary and secondary production in inshore environments are enhanced by organic matter and nutrients derived from river discharge, terrestrial/anthropogenic emissions, and the resuspension of sediments (Bauer et al. 2013; Wang et al. 2019b). Further, these resources often are exported to the open ocean, forming an inshore-to-offshore continuum, which may show environmental gradients in primary production, nutrients, salinity, etc. (Bauer et al. 2013). Recent investigations have supported the classic biogeographic patterns in microbial diversity (such as distance-decay and species-area relationships) across environmental gradients of coastal waters and the factors that shape these patterns (Chen et al. 2023; Wang et al. 2015; Zhang et al. 2018; Zinger et al. 2014). However, regional variability and specific biogeographic patterns of microbial association (co-occurrence) features either within each domain of life or across the three domains along coastal environmental gradients remain elusive.

Network analyses have been extensively used to describe microbial co-occurrence relationships (Fath et al. 2007; Faust and Raes 2012; Kishore et al. 2023). Although the relationship between statistically inferred co-occurrence and exact ecological interactions is still poorly understood (Blanchet et al. 2020; Kishore et al. 2023), inferring microbial co-occurrence networks from large microbiome sequencing datasets is a powerful tool to explore hidden patterns of microbial consortia that cannot be directly observed (Fath et al. 2007; Kishore et al. 2023; Liu et al. 2023). By unveiling ecologically meaningful associations and identifying hub/keystone species, co-occurrence networks have the potential to help infer the functional roles or environmental niches occupied by uncultured microorganisms (Barberán et al. 2012). Some previously undiscovered microbial co-occurrence relationships in marine environments have been detected by employing network analysis to microbiota datasets (Needham and Fuhrman 2016; Parada and Fuhrman 2017; Reji et al. 2019). However, efforts to depict marine microbial biogeographic patterns in light of co-occurrence relationships are rare. Therefore, we propose that integrating network features with microbial biogeography can aid understanding the significance of interspecific associations in shaping the distribution and assembly of the complex microbiome in coastal ecosystems.

The coastal area of northern Zhejiang Province, China, is an interconnected marine ecosystem in the East China Sea covering over 250-km and consisting of various ecological zones spanning from inshore to offshore. These include Hangzhou Bay, which is one of the most eutrophic bays in China that is highly influenced by the runoff from the Qiantang River and agricultural and industrial discharges; Xiangshan Bay and Sanmen Bay, which are two semi-enclosed bays having excess nutrient loads due to terrestrial emissions, extensive mariculture, and poor water exchange; Zhoushan Islands, located at the mouth of Hangzhou Bay, which are barriers from the estuary to the ocean; the coastal eastern Xiangshan, which contains fishery hot spots; Yushan Islands, which is a reserve area; and the boundary of the Island Chain, which comprises the east boundary of this coastal water area (Fig. 1) (Wang et al. 2015, 2019a). These diversified marine functional zones form a subtropical coastal continuum with multiple environmental gradients (Fig. S1), serving as a model to explore biogeographic patterns of microbial co-occurrence features across complex environmental gradients. In this study, using an integrated 16S/18S rRNA microbiome dataset from three domains at the regional scale, we aim to address three major questions: (1) How do the topological features of microbial co-occurrence networks vary across the three domains? (2) Is there a consistent biogeographic pattern of microbial co-occurrence networks features across the three domains? (3) Which taxa or environmental factors drive biogeographic shifts of co-occurring microbial assemblages across the environmental gradients?

Methods

Study area, sampling, and physicochemical analyses

Surface water samples were collected from 99 stations along the coastal waters of northern Zhejiang in the East

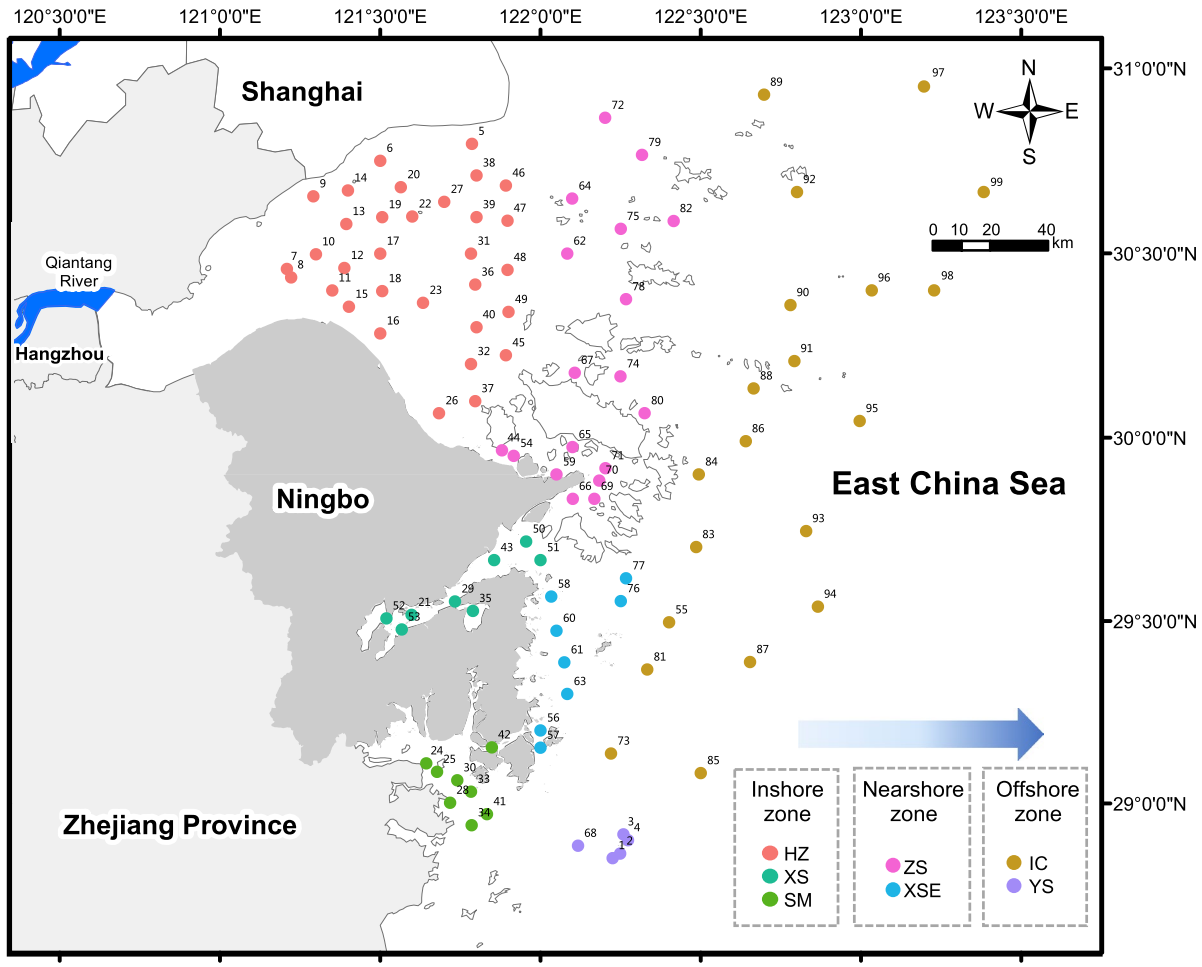

Fig. 1 Map of sampling stations. The 99 sampling stations were distributed in seven zones: three inshore zones including Hangzhou Bay (HZ, 32 stations), Xiangshan Bay (XS, 8 stations), Sanmen Bay (SM, 8 stations); two nearshore zones including the Zhoushan Islands (ZS, 18 stations), the eastern Xiangshan (XSE, 8 stations); and two offshore zones including the boundary of the Island Chain (IC, 20 stations) and Yushan Reserve (YS, 5 stations)

China Sea, during March 6–20, 2015. These stations are distributed in seven representative zones, including Hangzhou Bay (HZ), Xiangshan Bay (XS), Sanmen Bay (SM), Zhoushan Islands (ZS), the eastern Xiangshan (XSE), Yushan Reserve (YS), and the boundary of the Island Chain (IC; Fig. 1). To make comparisons along this coastal continuum, we further classified the seven zones into three major categories based on geographic proximity and similarity of environmental conditions. Three bays (HZ, XS and SM) near the mainland subject to the strongest terrestrial impacts were classified as 'inshore zones'. Two zones (YS and IC) located towards the open sea at greater distances from the shore (>45 km), in deeper water depths (>25 m), and subject to less terrestrial impacts were classified as 'offshore zones'. The other two geographically and environmentally intermediate zones (ZS and XSE) located between the inshore and offshore zones were classified as 'nearshore zones'.

Water samples were collected from the surface layer (0.5-m depth) using a 5-L Niskin sampler on the boat deck. Approximately 600 mL of seawater were pre-filtered using a 100-μm sterilized nylon filter, and then were filtered through a 0.2-μm polycarbonate membrane (Millipore, USA). The membranes were frozen at −20 °C on board and then at −80 °C after being transported to the laboratory. To prevent cross contamination, water samplers and filtration systems were washed with sterile water and thoroughly rinsed with the local water at each station before water collection and filtration. Water temperature, salinity, and pH of water samples were measured in situ. Chlorophyll-*a*, chemical oxygen demand (COD), dissolved oxygen (DO), total organic carbon (TOC), total nitrogen, total phosphorus, suspended

particles, phosphate, silicate, nitrate, nitrite and ammonium were measured using standard methods (AQSIQ 2007). The Eutrophication index (*EI*) was calculated as follows: $EI = \rho(\mathrm{COD}) \times \rho(\mathrm{DIP}) \times \rho(\mathrm{DIN}) \times 10^6/4500$, where ρ(COD), ρ(DIP), and ρ(DIN) are the concentrations of COD, phosphate, and inorganic nitrogen (the sum of nitrite, nitrate, and ammonium) (Jiang et al. 2015).

DNA extraction, PCR amplification, and Illumina sequencing

Total DNA from the membranes was extracted using the Power Soil DNA Isolation Kit (MOBIO, USA). An aliquot of 10 ng purified DNA template from each sample was amplified in triplicate with a 20-μL PCR reaction system. The hypervariable regions of the bacterial and archaeal 16S rRNA genes and eukaryotic 18S rRNA genes were amplified with the primer sets and PCR conditions in Table S1. An equimolar amount of PCR product for each sample was mixed into one pooled sample for each primer set and purified using the PCR fragment purification kit (TaKaRa, Japan). Sequencing libraries were generated using TruSeq™ DNA PCR-Free Library Preparation Kit, and the 2 × 250 bp paired-end sequencing was performed on the Illumina MiSeq platform (Illumina, USA).

Processing of sequence data

Paired reads of Bacteria, Archaea, and Microeukaryotes were joined using FLASH with default settings (Magoč and Salzberg 2011) and then processed using QIIME v1.9.1 (Caporaso et al. 2010), respectively. Briefly, the sequences were quality filtered at a Phred quality score (Q) < 20 using the script *split_libraries_fastq.py*. After detection and removal of chimeras using UCHIME (Edgar et al. 2011), Operational Taxonomic Units (OTUs) were clustered with 97% sequence similarity cutoff using UCLUST (Edgar 2010). Regarding the choice of methods for denoting microbial taxa, recent studies often found that the use of Exact Sequence Variants (ESVs, also known as amplicon sequence variants (ASVs) or Zero-radius OTUs (ZOTUs)) or OTUs (97% sequence similarity cutoff) showed similar ecological patterns across large field-based datasets (García-García et al. 2019; Glassman et al. 2018; Grady et al. 2019; Kerrigan and D'Hondt 2022). Since using ESVs may overestimate microbial richness, especially for Microeukaryotes that have high intraspecific variability in 18S rRNA gene copies (Caron and Hu 2019; Lavrinienko et al. 2021), which were often clustered at 97% sequence similarity (Aslani et al. 2022; David et al. 2021; Mo et al. 2021), we used OTUs for Microeukaryotes. For consistency and integration of the three domains, Bacteria and Archaea were also analyzed based on OTUs. The representative sequence of each OTU was taxonomically assigned using the SILVA 128 database (Quast et al. 2013). To further improve microbial taxonomic annotation, prokaryotic (mainly for Archaea) and microeukaryotic OTUs were complementarily assigned using the Genome Taxonomy Database (GTDB) (Parks et al. 2021) and the Protist Ribosomal Reference (PR^2) database (Guillou et al. 2012). Any OTU containing less than 0.0001% of the total sequences for each domain was removed to avoid potential spurious OTUs (Bokulich et al. 2013). Chloroplast, mitochondria and unassigned OTUs were removed, as were Metazoa from the microeukaryotic OTU table. After the above procedures, 2,722,454 (read range 19,105–39,458, mean = 22,499.5 per sample), 3,236,337 (read range 25,496–39,899, mean = 32,690.3 per sample), and 2,925,841 (read range 15,204–38,818, mean = 29,553.9 per sample) high-quality clean reads resulted for the bacterial, archaeal, and microeukaryotic OTUs tables, respectively. The three OTU tables were rarefied at 10,000 reads per sample for downstream analyses to normalize differences in sequencing depth. Rarefaction curves showed that sampling was sufficient to capture the trend of α-diversity in the samples among the three microbial domains (Fig. S2). The α-diversity for the three microbial domains was calculated using QIIME with multiple indices including Chao 1, observed species, phylogenetic diversity (PD) whole tree and the Shannon–Wiener index.

Microbial intra-domain network

The intra-domain co-occurrence networks were constructed separately for Bacteria, Archaea, and Microeukaryotes to visualize the interspecies associations within each domain. To increase network sensitivity, each OTU table was restricted to those OTUs that comprised > 100 total reads and were present in > 25% of samples (Needham and Fuhrman 2016). For each domain, OTUs correlations were calculated using SparCC based on 10 iterations and 100 bootstrap replications whilst accounting for their inherent sparsity and compositionality (Friedman and Alm 2012). Only significant correlations ($P < 0.01$) with SparCC $|r| > 0.6$ were used for the network using the R package 'igraph' (Csárdi and Nepusz 2006) to reduce spurious correlations. Networks were visualized using Gephi (Bastian et al. 2009) and the modules were detected using the Louvain algorithm (Blondel et al. 2008). Ternary plots were used to identify the geographic distribution of each network module in the inshore, nearshore, and offshore zones. A set of network-level topological features (as listed and defined in Table S2) was calculated as previously described (Newman 2003). Higher average degree, the clustering coefficient, graph density, and lower average path lengths suggest a more

Table 1 Comparisons among the topological features for the intra-domain co-occurrence networks of the three microbial domains

Topological features	Bacteria	Archaea	Microeukaryotes
Number of nodes	718	123	207
Number of edges	31433	1843	875
Positive edges	23889	1263	829
Negative edges	7544	580	46
Modularity	0.85	1.28	0.70
Graph density	0.122	0.246	0.041
Average degree	87.56	29.97	8.45
Betweenness centrality	464	57	205
Clustering coefficient	0.81	0.78	0.65
Average path length	2.32	1.94	3.49

connected network (Ma et al. 2016). Furthermore, we generated sub-networks for each station from three intra-domain networks and then calculated the topological features for each sub-network as described by Ma et al. (2016). The spatial distribution of station-based topological features was visualized using Kriging interpolation in ArcGIS 10.6. Scatter diagrams were used to analyze the relationships between topological features and environmental factors across the sampling zones.

Microbial inter-domain network

The inter-domain network was constructed to visualize the co-occurrence relationships of taxa across the three domains and their correlations with environmental factors. Here we used stricter criteria than intra-domain networks to narrow the scope of focus and to reduce any compositional effects (Weiss et al. 2016). The OTUs that comprise > 500 total reads in all samples and were present in > 25% of samples in each domain (accounting for 72.2%, 97.1%, and 91.1% of bacterial, archaeal, and microeukaryotic reads, respectively) were included in the SparCC calculations. Only significant correlations ($P < 0.01$) with SparCC $|r| > 0.75$ were visualized in the networks using Cytoscape 3.10.1. Node-level topological features (as listed and defined in Table S2) were further calculated using the NetworkAnalyzer plugin in Cytoscape (Assenov et al. 2008). Nodes with a core position in the network tend to have higher values for degree, betweenness centrality, clustering coefficient and lower values for average shortest path length (Ma et al. 2016). The generation of the inter-domain sub-networks and the calculation of the station-based topological features were performed in the same way as for the intra-domain sub-network. The *Zi-Pi* plot was used to distinguish the roles that each node played in the network by analyzing two parameters including within-module connectivity (*Zi*) and among module connectivity (*Pi*). The roles of nodes can be classified into four different categories, including peripherals ($Zi < 2.5$, $Pi < 0.62$), connectors ($Zi < 2.5$, $Pi > 0.62$), module hubs ($Zi > 2.5$, $Pi < 0.62$), and network hubs ($Zi > 2.5$, $Pi > 0.62$) (Guimerà and Nunes Amaral 2005; Olesen et al. 2007).

Results

Environmental gradients from inshore to offshore zones

Multiple environmental gradients along the inshore-nearshore-offshore continuum were visualized using boxplots (Fig. S1). Suspended particles concentrations showed the greatest variation among zones, ranging from 1,109 mg/L in the inshore zones to only 63 mg/L in the offshore zones. Nutrient-related parameters, including total nitrogen (0.77–1.31 mg N/L), nitrate (0.61–1.00 mg N/L), phosphate (0.026–0.046 mg P/L), silicate (1.01–1.53 mg Si/L), COD (0.92–1.77 mg/L), and TOC (1.45–1.59 mg/L), as well as DO (8.86–9.29 mg/L), showed a similar pattern to suspended particles, with a gradual decrease along the inshore-nearshore-offshore continuum. Furthermore, the spatial variation of eutrophication status in this coastal area was also assessed using a eutrophication index (*EI*), which integrates several nutrient parameters, including nitrate, nitrite, ammonium, phosphate, and COD. The results showed a clear decreasing trend of *EI* from the inshore to the offshore zones. Both the inshore ($EI = 22.11$) and nearshore ($EI = 9.67$) zones were heavily eutrophic ($EI > 9$), especially some stations located in HZ with *EI* values above 50. Conversely, natural factors, including water temperature (8.88–11.05 °C), salinity (20.76–29.13 psu), and pH (8.08–8.17) significantly increased from the inshore to the offshore zones.

An overview of the bacterial, archaeal, and microeukaryotic communities

We obtained 29,275 bacterial OTUs (ranging from 544 to 6,618 OTUs), 2,818 archaeal OTUs (ranging from 120 to 962 OTUs), and 9,027 microeukaryotic OTUs (ranging from 345 to 2,088 OTUs) from 99 seawater samples. The most abundant bacterial Phyla and proteobacterial Classes were the Gammaproteobacteria (average relative abundance of 37.4 ± 23.4%), Alphaproteobacteria (10.7 ± 7.0%), Betaproteobacteria (9.2 ± 6.0%), Actinobacteriota (6.8 ± 3.9%), Deltaproteobacteria (5.8 ± 4.0%), Chloroflexota (5.8 ± 4.5%), and Bacteroidota (5.2 ± 3.5%). Thaumarchaeota Marine Group I (MGI, now placed within the Family Nitrosopumilaceae of the Phylum Thermoproteota in the GTDB taxonomy, 92.1 ± 5.4%), primarily *Nitrosopumilus* (80.5 ± 8.3%), dominated the archaeal community.

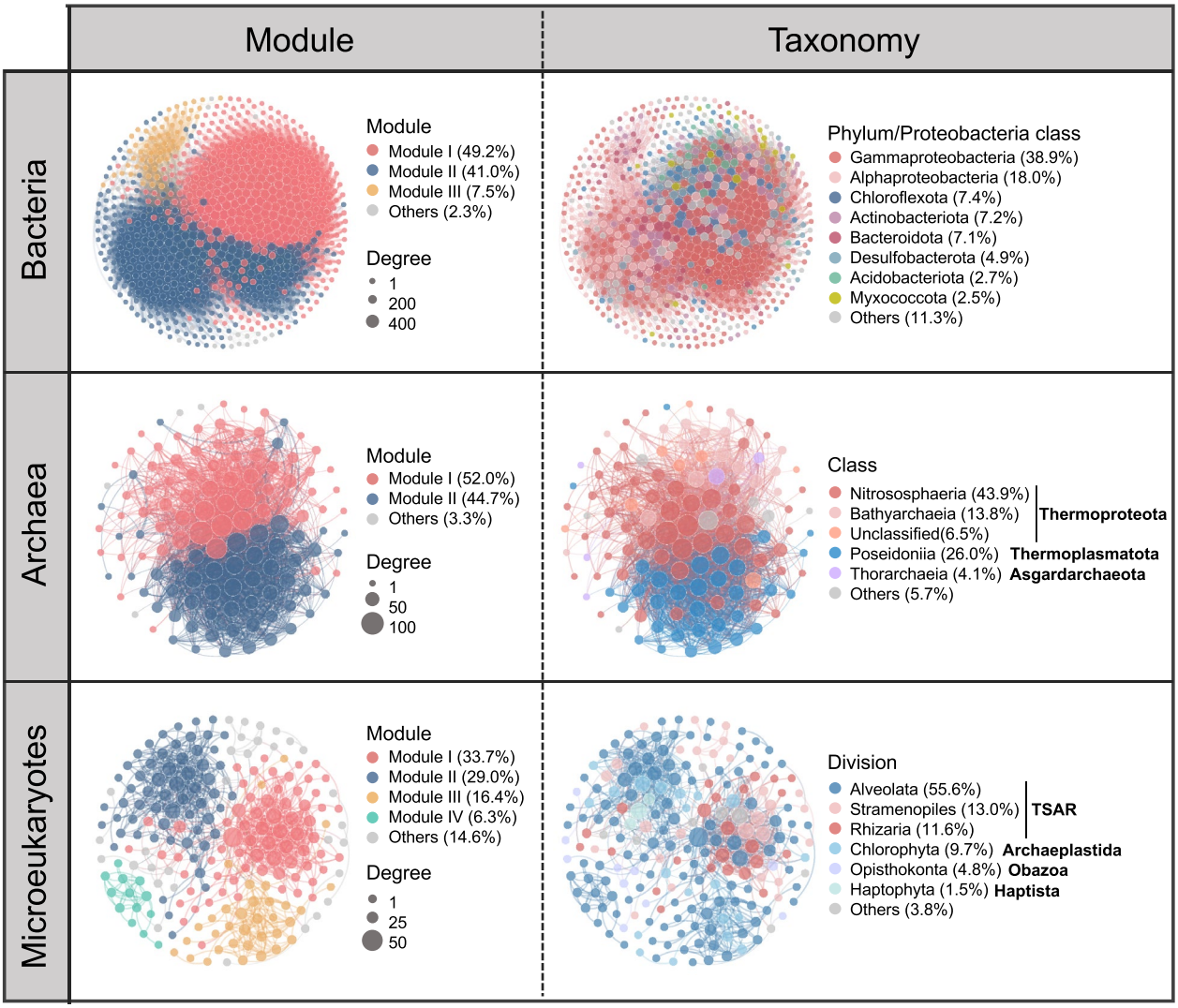

Fig. 2 Intra-domain networks showing co-occurrence patterns within each domain. The size of each node is proportional to its degree. Pairwise SparCC $|r|>0.6$ with $P<0.01$ is shown as edges of the networks. In the left panels, nodes are colored by module. In the right panels, nodes are colored according to taxonomy

Microeukaryotic communities were dominated by the TSAR Supergroup (88.2 ± 7.1%), comprising the Divisions Telonemia (0.3 ± 0.3%), Stramenopiles (6.6 ± 7.0%), Alveolata (71.4 ± 14.3%, mainly Dinoflagellata (40.5 ± 24.1%) and Ciliophora (27.3 ± 26.7%)), and Rhizaria (9.9 ± 8.2%).

Archaeal α-diversity (including Chao 1, observed species, PD, and Shannon–Wiener index) was the lowest among the three domains, while the α-diversity was the highest for the Bacteria (Fig. S3). The α-diversity in all three domains was always significantly lower in the offshore zones than in the inshore and nearshore zones. Moreover, the inshore and nearshore zones had similar microbial diversity by most indices, with a few exceptions (e.g. the Shannon–Wiener index for Archaea, PD and the Shannon–Wiener index for Microeukaryotes peaked in the nearshore zones).

Topological features of microbial intra-domain co-occurrence networks

Microbial intra-domain networks were constructed (Fig. 2) and the network-level topological features (Table S2) were calculated to examine taxa correlations within each domain. Intra-domain networks captured 31,433, 1,843, and 875 edges among 718, 123, and 207 nodes for Bacteria, Archaea, and Microeukaryotes, respectively (Fig. 2, Table 1). The bacterial network was the most complex among the three domains,

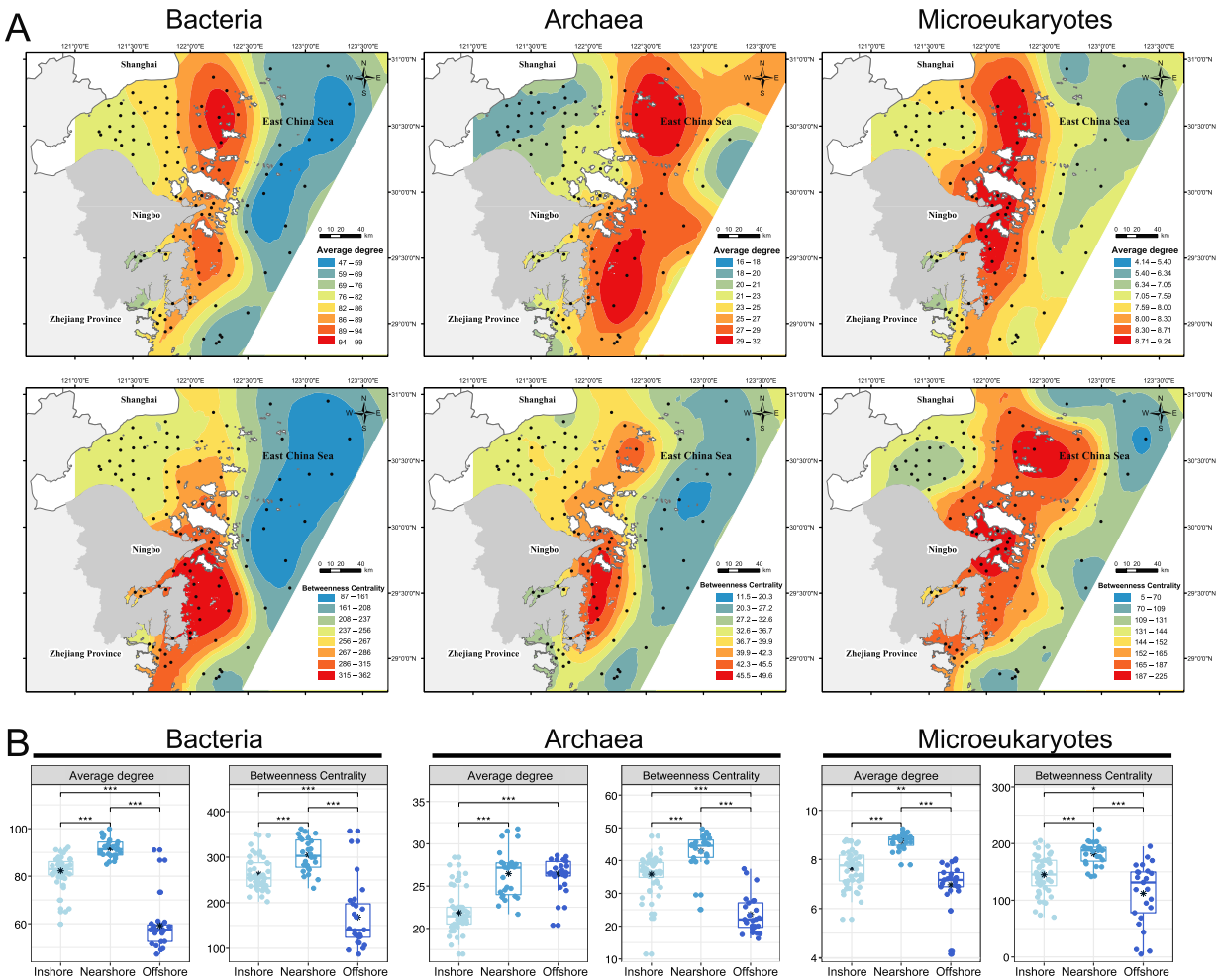

Fig. 3 Geographic patterns of topological features (average degree and betweenness centrality) of intra-domain sub-networks for the three microbial domains. **A** Kriging maps visualizing the spatial distribution of degree and betweenness centrality across sampling zones. **B** Boxplots showing the differences of these two topological features in the inshore, nearshore and offshore zones. The lower and upper hinges of the boxes correspond to the 25th and 75th percentiles. The lines and asterisks (*) in the boxes correspond to median and mean. The upper whisker extends from the hinge to the value no further than 1.5IQR (inter-quartile range). The lower whisker extends from the hinge to the smallest value at most 1.5IQR. Data beyond the end of the whiskers are "outlying" points (multiple comparisons after Kruskal–Wallis test, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$)

having the highest values for node number, average degree, betweenness centrality, and clustering coefficient (Table 1). The microeukaryotic network had the lowest values for density, average degree, betweenness centrality, and clustering coefficient as well as the highest value for average path length. Although the archaeal network had the minimum node number, its edge number and average degree were 2.11- and 3.55-fold higher than in the Microeukaryotes.

The station-based topological features (e.g. average degree, betweenness centrality, numbers of node and edge) of the intra-domain sub-networks for the three domains showed similar geographic patterns. Significantly higher values of topological features were observed for the nearshore zones (XSE and most ZS stations) than for the inshore (HZ, XS, and SM) and offshore zones (YS and IC) (Figs. 3 and S4). Furthermore, the values of average degree and betweenness centrality of the intra-domain sub-networks generally peaked at intermediate levels for the environmental factors, including water temperature, salinity, pH, DO, and nutrient-related parameters (including suspended particles, silicate, total nitrogen, and *EI*; Fig. S5).

Modularity of microbial intra-domain co-occurrence networks

Modularity analysis showed that the bacterial co-occurrence network could be divided into three primary

modules (Fig. 2). The majority of Module I were members of Gammaproteobacteria (node proportion 29.3%), Chloroflexota (13.9%), and Alphaproteobacteria (11.4%); Module II was mainly composed of Gammaproteobacteria (52.1%) and Alphaproteobacteria (21.9%); Members of Alphaproteobacteria (43.3%), Gammaproteobacteria (21.7%), and Bacteroidota (21.7%) constituted Module III. The archaeal network had the highest modularity of the three domains (Table 1). The two main modules were dominated by Nitrososphaeria (43.8% in Module I) and Poseidoniia (54.6% in Module II), respectively (Fig. 2). The microeukaryotic network had the lowest modularity (Table 1) and was clustered into four primary modules (Fig. 2). All four of those modules were dominated by Alveolata (ranging from 40.6 to 100%). In addition, Module I was co-dominated by Rhizaria (29.0%) and Stramenopiles (17.4%). Chlorophyta was the dominant member of Module II (16.7%) and Module III (20.6%) after Alveolata.

Module I of the co-occurrence networks for the three domains were distributed in both the inshore and nearshore zones (Fig. S6). Module II tended to be distributed in the offshore zone, although some taxa were shared among the three zones. Module III of the microeukaryotic network was almost evenly distributed among three zones, while its Module IV and bacterial Module III occurred in the offshore and inshore zones, respectively.

Topological features of microbial inter-domain co-occurrence network

The microbial inter-domain network had 198 nodes (including 139 bacterial OTUs, 35 archaeal OTUs, and 24 microeukaryotic OTUs), linked by 1,103 edges (including 627 positive edges and 476 negative edges; Fig. 4). The nodes belonging to the Bacteria and Archaea and the edges between these two domains accounted for 87.9% and 86.4% of the entire network, respectively (Fig. 4A), illustrating the predominant role of bacterial-archaeal associations in the microbial inter-domain network. Moreover, the node-level topological features (e.g. degree, betweenness centrality) for the archaeal nodes were significantly greater than those for the bacterial or microeukaryotic nodes (Fig. 4D).

In addition, the station-based sub-networks generated from the inter-domain network showed a similar geographic pattern in topological features (including average degree, betweenness centrality, and numbers of node and edge; Fig. S7) as those of the intra-domain sub-networks (Figs. 3 and S4). That is, the nearshore zones had significantly higher values of the topological features above than the inshore and offshore zones (Fig. S7).

Modularity of microbial inter-domain co-occurrence network

Modularity analysis revealed two primary modules in the inter-domain microbial co-occurrence network (Fig. 4B). Module A consisted of 78 nodes and 207 edges (positive: 176 vs. negative: 31), and Module B had 72 nodes and 247 edges (positive: 237 vs. negative: 10). Both modules exhibited clear geographic patterns (Figs. 4C and S8). Approximately 70% of Module-A-OTUs were distributed in the inshore and nearshore zones (referred as Inshore zones enriched and Inshore & Nearshore shared), while more than 78% of Module-B-OTUs were solely distributed in the offshore zones (referred as Offshore zones enriched).

The topological roles of nodes in the inter-domain co-occurrence network are shown in the *Zi-Pi* plot (Fig. 4E). All nodes identified as hubs (including 3 network hubs and 5 module hubs) are archaeal OTUs, mainly belonging to MGI and the Euryarchaeota Marine Group II (MGII, now placed in the Order Poseidoniales of the Phylum Thermoplasmatota in the GTDB taxonomy). Two MGII archaeal OTUs (*Poseidonia* [Arc_11] and MGIIb-O1 [Arc_21]) and one MGI OTU (*Nitrosopumilus* [Arc_14]) were classified as network hubs in our analysis. Two groups of MGI OTUs (that we defined as MGI_hubs_A and MGI_hubs_B) were respectively identified as module

(See figure on next page.)

Fig. 4 Co-occurrence patterns and ecological roles of three microbial communities in inter-domain network across the study area. Inter-domain co-occurrence network with nodes are colored according to three microbial domains (**A**), modules (**B**), and zones (**C**), respectively. The connection indicates a strong (SparCC $|r|>0.75$) and significant ($P<0.01$) correlation. The width of the lines indicates correlation strength. The size of each node is proportional to its degree. A summary of node (or node-edge) statistics is provided at the bottom of each network. Specifically, **A** numbers in the solid circles represent the number of nodes belonging to the corresponding microbial domains and the numbers adjacent to edge connections represent the number of cross-domain associations; **B** a doughnut chart shows the proportion of each module in the entire network, and the numbers represent the number of nodes belonging to the corresponding modules; **C** the bar chart shows the proportion of nodes belonging to the corresponding category in the entire network and each module according to nodes categories in Figure S8. **D** Node-level topological features associated with different microbial domains (multiple comparisons after Kruskal–Wallis test, $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$). **E** *Zi-Pi* plot showing the module-based topological role of each node in the inter-domain network. The threshold values of *Zi* (within-module connectivity) and *Pi* (among-module connectivity) for categorizing were 2.5 and 0.62, respectively. The taxonomic information of network hubs, module hubs and connectors are listed in Table S3

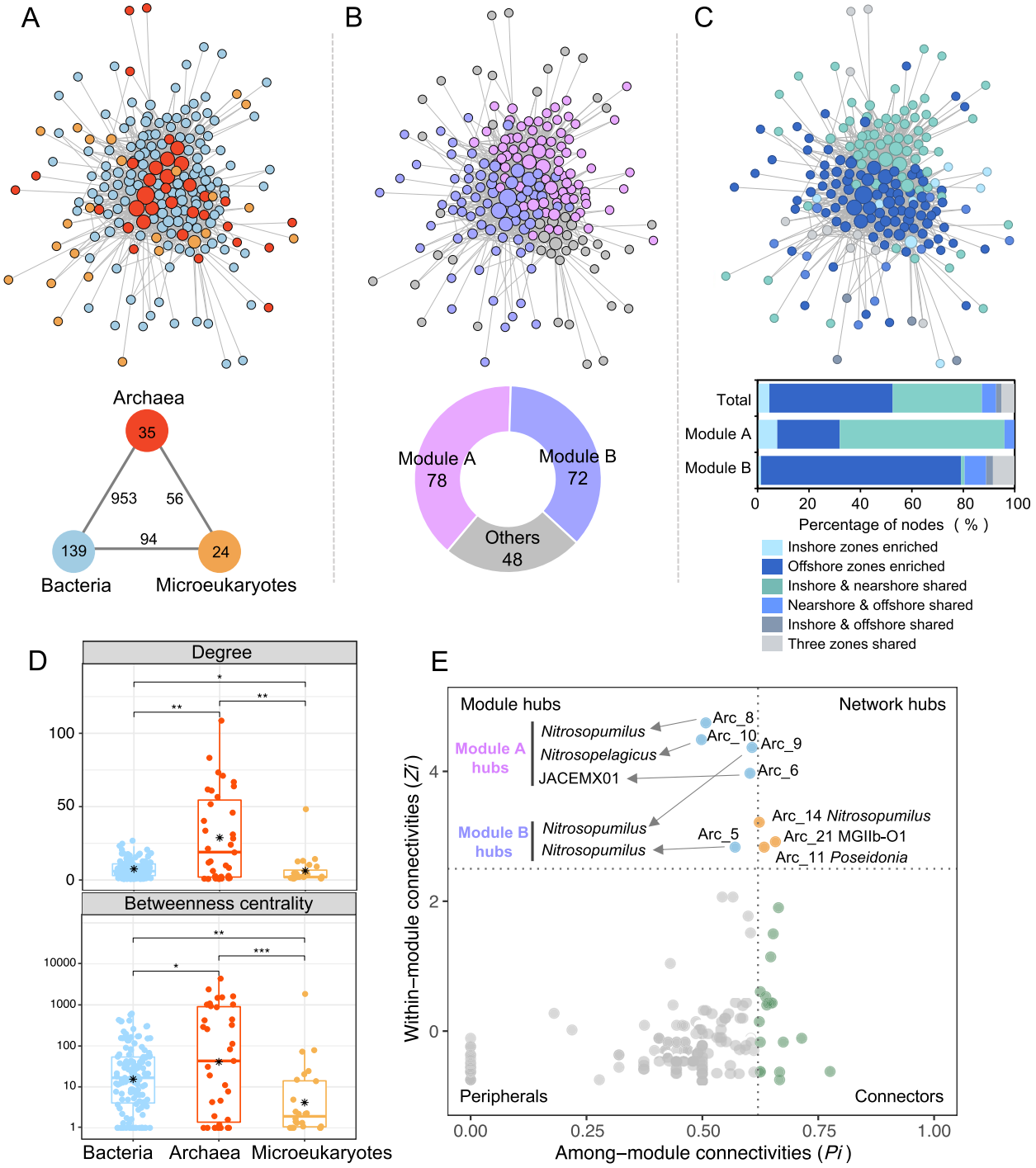

Fig. 4 (See legend on previous page.)

hubs in two network modules (Figs. 4E, 5, and Table S3). The MGI_hubs_A members (JACEMX01 [Arc_6], *Nitrosopumilus* [Arc_8], and *Nitrosopelagicus* [Arc_10]) showed positive correlations with bacterial OTUs mainly containing Woeseiaceae, Nitrosomonadaceae, Nitrospiraceae, and Trueperaceae (Fig. 5A). Additionally, some bacterial OTUs (members of Halieaceae, Rhodobacteraceae, and Alteromonadaceae), microeukaryotic OTUs (members of Prymnesiaceae), and archaeal MGIIb OTUs showed negative relationships with MGI_hubs_A.

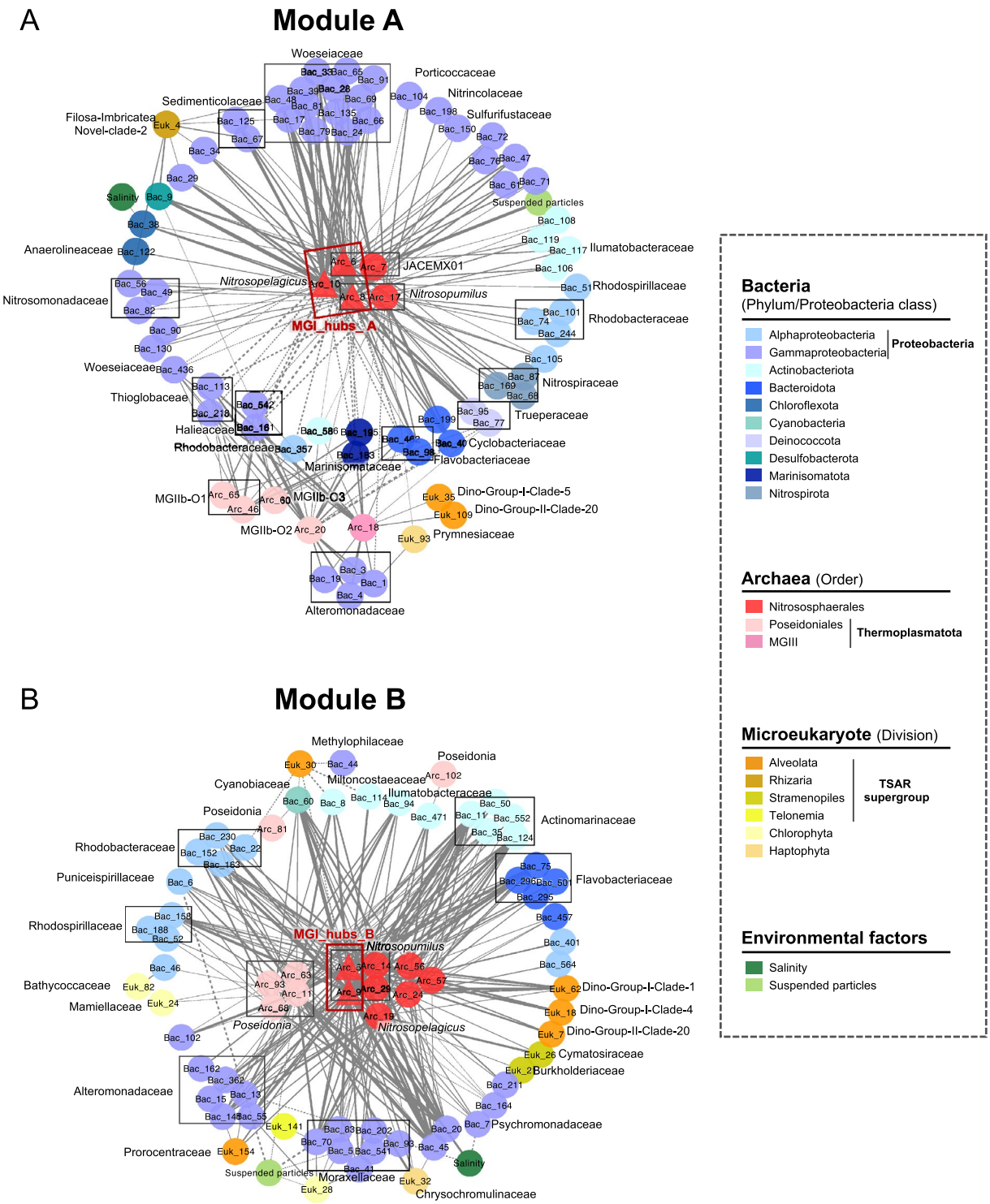

Fig. 5 The co-occurrence relationships of the three-domain microbes in Module A (**A**) and Module B (**B**) detected in the inter-domain network in Fig. 4B. Boxes enclose OTUs of major groups that appeared affiliated to the same bacterial and microeukaryotic family or archaeal genus. Triangular nodes represent module hubs. The width of the lines indicates correlation strength, and the solid and dashed lines indicate positive and negative correlations, respectively

Compared with MGI_hubs_A, members of MGI_hubs_B (*Nitrosopumilus* [Arc_5 and Arc_9]) had positive correlations with a greater number of microeukaryotic OTUs (mainly affiliated to Dino-Group-I and II of the Class Syndiniales; Fig. 5B). In addition, the MGI_hubs_B, together with MGII OTUs (members of *Poseidonia*), were positively correlated with bacterial OTUs, mainly from Alteromonadaceae, Moraxellaceae, Actinomarinaceae, Rhodobacteraceae, Rhodospirillaceae, and Flavobacteriaceae.

Correlations between microorganisms and environmental factors (mainly suspended particles and salinity) were seen in both modules (Fig. 5). Suspended particles and salinity in Module A were significantly positively correlated with a MGI archaeal OTU (JACEMX01 [Arc_7]) and negatively correlated with a bacterial OTU (Chloroflexota [Bac_38]), respectively (Fig. 5A). These two environmental factors were significantly correlated with a greater number of microbial taxa in Module B (Fig. 5B). Salinity was negatively correlated with two bacterial OTUs (Psychromonadaceae [Bac_7] and Alteromonadaceae [Bac_13]) and positively correlated with a MGI archaeal OTU (*Nitrosopumilus* [Arc_9]), while suspended particles were negatively correlated with multiple OTUs from three domains, including members of Moraxellaceae [Bac_5 and Bac_70], Puniceispirillaceae [Bac_6], *Poseidonia* [Arc_11], and Chlorellales [Euk_28] (Fig. 5B).

Discussion

Topological features of microbial co-occurrence networks varied among the three domains

Network analysis can potentially provide valuable perspectives on complex microbial communities beyond diversity and community composition by evaluating interspecific linkages within and among the three microbial domains. Our research found the highest α-diversity for the bacterial community among the three domains, as well as the highest complexity and connectivity for the bacterial intra-domain network, suggesting that the density of network connections might be affected by the diversity of interconnected OTUs, as was previously reported for a coastal water microbial network (Reji et al. 2019). Archaea had the lowest α-diversity and minimum network nodes among the three domains. However, Archaea's higher network connectivity and complexity than the Microeukaryotes indicated that some other factors besides diversity affected their network's structure.

Modularity is another key structural feature of complex networks (Newman 2006). Identification of modules has revealed not only that structured organization can be tightly linked, but perhaps more importantly, a link between this organization and function and stability (Lurgi et al. 2019). In our study, both intra-domain and inter-domain co-occurrence networks were partitioned into distinct modules, potentially representing shared ecological/functional niches or common metabolic interactions among taxa (Reji et al. 2019). The archaeal co-occurrence network showed the highest modularity among the three domains, which might suggest cooperative behavior within Archaea and the development of diverse mechanisms for archaeal colonization under a wide variety of environmental conditions (Gianetto and Heydari 2015; Takemoto and Borjigin 2011). This appeared coincident with high environmental adaptability and broad niche breadth of Archaea that we previously reported in coastal waters of southern Zhejiang Province, China (Wang et al. 2020). It was also found that each module of the co-occurrence networks contained members from the same or a restricted number of Phyla/Class/Division(s), suggesting possible taxonomic specificity for microbial module assembly. In addition, the distribution of all modules in this study revealed geographic patterns. For instance, the primary module for the bacterial co-occurrence network across the inshore-nearshore-offshore continuum showed a shift from Module III to Module I, and then to Module II. The taxonomic specificity and geographic patterning of the network modules suggest that phylogenetically related microorganisms may co-occur, possibly as a result of niche overlap among them (Cheung et al. 2018; Friedman and Alm 2012).

A consistent biogeographic pattern of microbial co-occurrence features for the three domains

Network topological features (including average degree, betweenness centrality, numbers of node and edge) for the three domains showed a consistent biogeographic pattern. At two ends of the environmental gradients, the inshore zones (especially HZ) and offshore zones had relatively weaker microbial associations, while the nearshore zones with intermediate terrestrial impacts (as indicated by multiple environmental factors including water temperature, salinity, pH, DO, and nutrient-related parameters) had stronger microbial associations (Fig. 6). Connell proposed an ecological theory termed the "intermediate disturbance hypothesis (IDH)" in 1978, which conceptualized the relationship between biodiversity and different levels of habitat disturbance, predicting that species richness is highest at an intermediate level of disturbance (Connell 1978). The consistently observed relationship between network topological features and terrestrial impacts, regardless of three microbial domains, appeared similar to the IDH theory. Thus, a modified IDH describing the relationship between the strength of microbial associations and different levels of

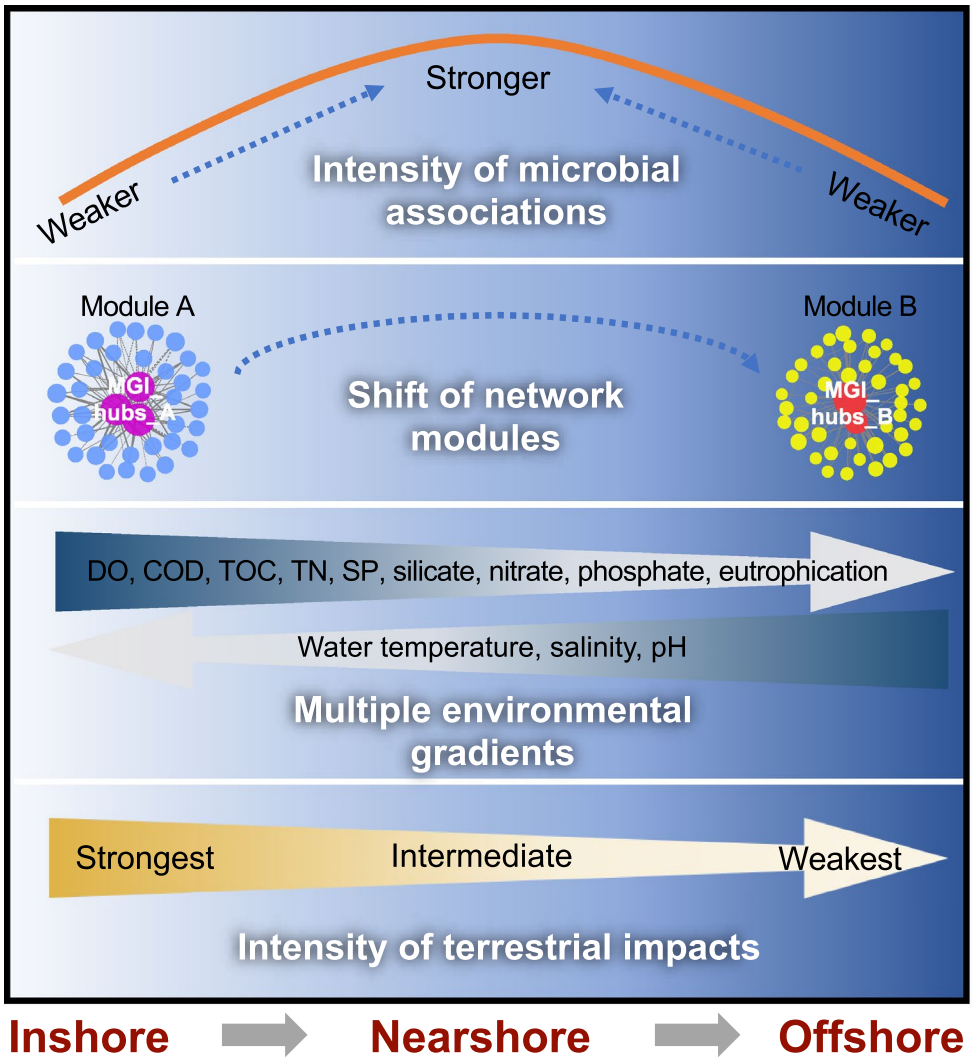

Fig. 6 Conceptual diagram showing the geographic patterns of microbial co-occurrence relationships across complex environmental gradients in subtropical coastal waters. Across a scale of about 250-km, as the intensity of terrestrial impacts decreased from inshore to offshore zones, dissolved oxygen (DO), chemical oxygen demand (COD), total organic carbon (TOC), total nitrogen (TN), suspended particles (SP), silicate, nitrate, phosphate, and eutrophication showed a gradually decreased gradient, while water temperature, salinity, and pH exhibited an opposite trend. Three microbial domains exhibited consistent network topology-based geographic patterns along environmental gradients. The inshore zones and offshore zones showed relatively weaker microbial associations, while the nearshore zones with intermediate terrestrial impacts had stronger microbial associations. Additionally, archaeal MGI OTUs were hubs for structuring inter-domain network, and the network had a structural shift from Module A to Module B along environmental gradients from inshore to offshore zones

terrestrial disturbance can be proposed as an open question for future verification.

In the present study, the offshore zones having minimal terrestrial impacts showed the weakest microbial associations, which could be largely determined by their lower microbial α-diversity compared with the other zones. However, high microbial α-diversity does not necessarily mean strong microbial associations. Notably, in the inshore zone HZ, we found the highest prokaryotic diversity, but relatively weak microbial associations. This zone has been very eutrophic for a long time, with its seawater quality worse than Grade IV (the worst grade; Bulletin of Marine Ecology and Environment Status of China in 2015) according to the Seawater Quality Standard of China (GB 3097–1997 and HJ 442–2008). Such eutrophic environments provide diverse inorganic nutrients and organic materials, and the abundant suspended particles also comprise diversified micro-spatial and resource niches for microorganisms. These may mitigate competition for nutrients and ecological niches,

leading to weaker associations among the microorganisms (Faust and Raes 2012; Ma et al. 2016). Compared to the inshore zones with the most intense terrestrial impacts and the offshore zones with the lowest microbial diversity, the strongest microbial associations were found in the nearshore zones with intermediate levels of terrestrial impacts. The trade-offs among community diversity, environmental conditions, ecological niches, and species interactions may contribute to this biogeographic pattern of microbial associations. The validity and universality of this pattern across various biomes and ecosystems are worthy of future testing.

Archaeal taxa were the hubs in the biogeographic shift between inter-domain network modules

Bacteria have long been considered the key players of the marine microbial food web owing to their sheer abundance and activity (Bunse and Pinhassi 2017). Our research confirmed that Bacteria had the highest proportion of nodes in inter-domain network, as well as the greatest intra-domain network complexity and connectivity, all of which could help stabilize community in the face of external disturbances (Dunne et al. 2002; Tilman et al. 1997). Compared to Bacteria, our knowledge of archaeal biology is limited. Nevertheless, important findings about Archaea in recent decades, such as the discovery of anaerobic methane oxidation (Boetius et al. 2000), thaumarchaeal ammonia oxidation (Leininger et al. 2006), globally distributed metabolic generalists Bathyarchaeota (Zhou et al. 2018), and the evolutionarily important Asgardarchaeota (Eme et al. 2023; Spang et al. 2015), are revolutionizing our understanding of Archaea and implying their importance as key interactive components in complex microbiomes (Moissl-Eichinger et al. 2018). Notably in this study, the nodes for Archaea rather than Bacteria had the highest degree and betweenness centrality in the inter-domain network, and archaeal OTUs (mainly MGI) were identified as module hubs or even network hubs, indicating that Archaea occupy a central position in the network.

MGI is a major archaeal group across the marine (particularly coastal) water column and its members are the predominant ammonia-oxidizers in the marine nitrogen cycle (Francis et al. 2005; Hugoni et al. 2013; Wuchter et al. 2006). Previous studies found that MGI appears to have distinct phylotypes segregated in different oceanic provinces (Massana et al. 2000), depths (Reji et al. 2019), or seasons (Parada and Fuhrman 2017), which could be associated with their differentiated physiological characters (Luo et al. 2014), life history strategies (Hugoni et al. 2013), and/or microbial co-occurrence patterns (Reji et al. 2019). In our survey, the MGI OTUs fell into two different groups (MGI_hubs_A and MGI_hubs_B), which were identified as the hub species of the two modules (Module A and Module B) in the inter-domain network. Along the inshore-nearshore-offshore continuum, the microbial inter-domain co-occurrence relationships shifted from Module A-dominated to Module B-dominated (Fig. 6), suggesting divergent biogeographic patterns for the microbial assemblages centered on the hub species in the two modules.

The MGI_hubs_A group in Module A had positive correlations with Nitrosomonadaceae and Nitrospiraceae OTUs, which are known as ammonia-oxidizing bacteria and nitrite-oxidizing bacteria, respectively (Kuypers et al. 2018). The co-occurrence of these taxa in Module A may be related to the fact that high nitrogen loads in the inshore and nearshore zones provided sufficient substrate for these nitrifiers, which have the potential to be partners for complete nitrification. The MGI_hubs_A also co-occurred with a group of OTUs affiliated with Woeseiaceae (formerly classified in the JTB255 marine benthic group), which are ubiquitous microbial members in diverse marine benthic environments, with hypothesized roles in carbon and sulfur metabolism (Bowman et al. 2005; Dyksma et al. 2016; Mußmann et al. 2017). Their occurrence in the surface waters of this study might be due to the resuspension of the coastal sediments. Trueperaceae (in the Phylum of Deinococcota) in Module A is likely involved in the degradation of organic matter (Qian et al. 2017). Therefore, the co-occurrence of MGI_hubs_A and the members of the above microbial groups having versatile metabolic roles in processing organic materials and inorganic nutrients, might interact to mediate the element cycles in nutrient-rich inshore and nearshore zones.

The MGI_hubs_B OTUs in Module B commonly co-occurred with the archaeal OTUs affiliated with the Poseidonaceae (aka. MGIIa) Genus *Poseidonia*. MGII is an abundant planktonic archaeal group in global oceans (Rinke et al. 2019), with indications of a photoheterotrophic lifestyle that combines proteorhodopsin-based phototrophy with the remineralization of high-molecular-weight organic matter (Tully 2019). In addition to MGII members, we also found some other presumed heterotrophs (e.g., Rhodobacteraceae, Actinomarinaceae, and Flavobacteriaceae) co-occurring in Module B, for which a rhodopsin-based photoheterotrophic lifestyle was indicated in previous investigations (Torre et al. 2003; Newton et al. 2010; Yoshizawa et al. 2014). In addition, more phototrophic microeukaryotes (e.g., Chlorellales, Prorocentraceae, and Cymatosiraceae) occurred in Module B. Some of these phototrophic taxa (e.g., Chlorellales, *Poseidonia*) were significantly negatively correlated with suspended particles (Fig. 5B). These results suggest that high light intensity due to more light penetration in

the offshore zones having fewer suspended particles was one of the potential drivers of the associations among these phototrophic assemblages. Collectively, our work suggests that Archaea, particularly MGI taxa, serve as hubs in inter-domain microbial co-occurrence networks in coastal waters, potentially mediating biogeographic shifts of microbial assemblages from inshore to offshore zones. Nevertheless, it remains uncertain whether the hub species identified by the network algorithms act as keystones in actual ecological communities (Faust 2021). Further experimental evidence, such as experiments that involve the removal/addition of hub species and randomly selected control species in a simplified synthetic community (Berry and Widder 2014), should be conducted to evaluate the roles of the inferred hub species.

Conclusions

Here we depicted microbial biogeography in terms of complex co-occurrence relationships by conducting high-coverage sampling across multiple environmental gradients of subtropical coastal waters, and integrated a dataset covering Bacteria, Archaea and Microeukaryotes. Network analysis for inferring co-occurrence patterns from large environmental datasets provides a new perspective for microbial biogeography beyond diversity and composition. However, we note that network-based microbial co-occurrences should not be uncritically assigned to biological interactions. Comparing co-occurrence networks and real-world interactions remains a major unresolved challenge, which requires the development of more accurate network inference algorithms and further experimental validation of predicted co-occurrence relationships.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13717-024-00550-4.

Supplementary material: Figure S1. Boxplots showing the changes of environmental factors from inshore to offshore zones. **Figure S2.** Rarefaction curves showing α-diversity indices (including Chao 1, observed species, phylogenetic diversity (PD), and the Shannon–Wiener index) of each sample at different sequencing depths. **Figure S3.** Boxplots showing the changes of α-diversity indices (including Chao 1, observed species, phylogenetic diversity (PD), and Shannon–Wiener index) of Bacteria, Archaea, and Microeukaryotes from inshore to offshore zones. **Figure S4.** Geographic patterns of topological features (the number of nodes and edges) for intra-domain sub-networks of three microbial domains. **Figure S5.** Scatter diagrams showing the relationships between topological features (average degree and betweenness centrality) of co-occurrence networks for three domains and environmental factors across the sampling zones. **Figure S6.** Ternary plots showing the distribution patterns of each module in the intra-domain co-occurrence networks for Bacteria, Archaea, and Microeukaryotes in different zones (inshore, nearshore and offshore). **Figure S7.** Boxplots showing the differences in topological features (the number of nodes and edges, degree, and betweenness centrality) of inter-domain sub-networks for three microbial communities in the inshore, nearshore and offshore zones. **Figure S8.** Ternary plots showing the distribution patterns of all OTUs (B) in the inter-domain co-occurrence network and the OTUs in Module A (C) and Module B (D) in different zones (inshore, nearshore and offshore). **Table S1.** Information for primers used in this study. **Table S2.** Description of topological features. **Table S3.** Detailed information about the network hubs, module hubs and connectors for the inter-domain co-occurrence network of three microbial communities.

Acknowledgements

The authors thank the Marine Environmental Monitoring Center of Ningbo, SOA, China, for providing ship time and assistance for collecting and analyzing the environmental profiles, and thank Peng Qian, Heping Chen, and Hanjing Hu for their help in sample collection, bench work, and data analysis. The authors thank Associate Editor Carol Stepien for many valuable suggestions and English corrections to improve the manuscript.

Author contributions

D.Z. and K.W. designed and organized the study; H.Z., K.W., and D.Z. performed the study; K.W. and D.H. strategized the data analysis; D.H., H.Y., and H.L. analyzed the data; D.H. and K.W. wrote the manuscript. All authors reviewed and approved the final manuscript for publication.

Funding

This work was supported by the National Natural Science Foundation of China (41977192), Zhejiang Provincial Natural Science Foundation of China (LY21D060004), Natural Science Foundation of Ningbo (2021J060 and 2019A610449), Fundamental Research Funds for the Provincial Universities of Zhejiang (SJLY2022001), and K.C. Wong Magna Fund in Ningbo University.

Availability of data and materials

The sequence data are available under accession numbers DRA010147 (Bacteria), DRA010148 (Archaea), and DRA010149 (Microeukaryotes) in the Sequence Read Archive of DDBJ (http://ddbj.nig.ac.jp/DRASearch).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Key Laboratory of Aquaculture Biotechnology of the Ministry of Education, School of Marine Sciences, Ningbo University, Ningbo, China. [2]Collaborative Innovation Center for Zhejiang Marine High-Efficiency and Healthy Aquaculture, Ningbo, China.

Received: 18 April 2024 Accepted: 9 September 2024
Published online: 03 October 2024

References

Aslani F, Geisen S, Ning D, Tedersoo L, Bahram M (2022) Towards revealing the global diversity and community assembly of soil eukaryotes. Ecol Lett 25(1):65–76

Assenov Y, Ramirez F, Schelhorn S-E, Lengauer T, Albrecht M (2008) Computing topological parameters of biological networks. Bioinformatics 24(2):282–284

AQSIQ (2007) The specification for marine monitoring of China–Part 4: seawater analysis (GB 17378.4–2007). General Administration of Quality Supervision, Inspection and Quarantine (AQSIQ) of the People's Republic of China

Barberán A, Bates ST, Casamayor EO, Fierer N (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J 6(2):343–351
Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media
Bauer JE, Cai W, Raymond PA, Bianchi TS, Hopkinson CS, Regnier PAG (2013) The changing carbon cycle of the coastal ocean. Nature 504(7478):61–70
Berry D, Widder S (2014) Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Front Microbiol 5:00219
Blanchet FG, Cazelles K, Gravel D (2020) Co-occurrence is not evidence of ecological interactions. Ecol Lett 23(7):1050–1063
Blondel V, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech-Theory Exp 2008:10008
Boetius A, Ravenschlag K, Schubert C, Rickert D, Widdel F, Gieseke A et al (2000) A marine microbial consortium apparently mediating anaerobic oxidation of methane. Nature 407(6804):623–626
Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R et al (2013) Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat Meth 10:57–59
Bowman JP, McCammon SA, Dann AL (2005) Biogeographic and quantitative analyses of abundant uncultivated γ-proteobacterial clades from marine sediment. Microb Ecol 49(3):451–460
Bunse C, Pinhassi J (2017) Marine bacterioplankton seasonal succession dynamics. Trends Microbiol 25(6):494–505
Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Meth 7:335–336
Caron DA, Hu SK (2019) Are we overestimating protistan diversity in nature? Trends Microbiol 27(3):197–205
Chen W, Sang S, Shao L, Li Y, Li T, Gan L et al (2023) Biogeographic patterns and community assembly processes of bacterioplankton and potential pathogens in subtropical estuaries in China. Microbiol Spectr 11:e03683-22
Cheung MK, Wong CK, Chu KH, Kwan HS (2018) Community structure, dynamics and interactions of bacteria, archaea and fungi in subtropical coastal wetland sediments. Sci Rep 8:14397
Connell JH (1978) Diversity in tropical rain forests and coral reefs. Science 199(4335):1302–1310
Csárdi G, Nepusz T (2006) The igraph software package for complex network research. InterJ Complex Syst 1695(5):1–9
David GM, López-García P, Moreira D, Alric B, Deschamps P, Bertolino P et al (2021) Small freshwater ecosystems with dissimilar microbial communities exhibit similar temporal patterns. Mol Ecol 30(9):2162–2177
de la Torre JR, Christianson LM, Béjà O, Suzuki MT, Karl DM, Heidelberg J et al (2003) Proteorhodopsin genes are distributed among divergent marine bacterial taxa. Proc Natl Acad Sci USA 100(22):12830–12835
Deutschmann IM, Delage E, Giner CR, Sebastián M, Poulain J, Arístegui J et al (2024) Disentangling microbial networks across pelagic zones in the tropical and subtropical global ocean. Nat Commun 15:126
Dunne JA, Williams RJ, Martinez ND (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecol Lett 5(4):558–567
Dyksma S, Bischof K, Fuchs BM, Hoffmann K, Meier D, Meyerdierks A et al (2016) Ubiquitous Gammaproteobacteria dominate dark carbon fixation in coastal sediments. ISME J 10(8):1939–1953
Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460–2461
Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27(16):2194–2200
Eme L, Tamarit D, Caceres EF, Stairs CW, De Anda V, Schön ME et al (2023) Inference and reconstruction of the heimdallarchaeial ancestry of eukaryotes. Nature 618(7967):992–999
Falkowski P (2012) Ocean science: the power of plankton. Nature 483(7387):S17–S20
Fath BD, Scharler UM, Ulanowicz RE, Hannon B (2007) Ecological network analysis: network construction. Ecol Modell 208(1):49–55
Faust K (2021) Open challenges for microbial network construction and analysis. ISME J 15(11):3111–3118
Faust K, Raes J (2012) Microbial interactions: from networks to models. Nat Rev Microbiol 10:538–550
Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. Proc Natl Acad Sci USA 102(41):14683–14688
Friedman J, Alm EJ (2012) Inferring correlation networks from genomic survey data. PLoS Comp Biol 8(9):e1002687
Fuhrman JA, Cram JA, Needham DM (2015) Marine microbial community dynamics and their ecological interpretation. Nat Rev Microbiol 13:133–146
García-García N, Tamames J, Linz AM, Pedrós-Alió C, Puente-Sánchez F (2019) Microdiversity ensures the maintenance of functional microbial communities under changing environmental conditions. ISME J 13(12):2969–2983
Gianetto DA, Heydari B (2015) Network Modularity is essential for evolution of cooperation under uncertainty. Sci Rep 5:9340
Glassman SI, Martiny JBH, Tringe SG (2018) Broadscale ecological patterns are robust to use of exact sequence variants versus operational taxonomic units. mSphere 3(4):e00148
Grady KL, Sorensen JW, Stopnisek N, Guittar J, Shade A (2019) Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. Nat Commun 10:4135
Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L et al (2012) The Protist Ribosomal Reference database (PR^2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res 41(D1):D597–D604
Guimerà R, Nunes Amaral LA (2005) Functional cartography of complex metabolic networks. Nature 433(7028):895–900
Hugoni M, Taib N, Debroas D, Domaizon I, Dufournel IJ, Bronner G et al (2013) Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. Proc Natl Acad Sci USA 110(15):6004–6009
Jiang M, Chen H, Chen Q, Wu H, Chen P (2015) Wetland ecosystem integrity and its variation in an estuary using the EBLE index. Ecol Indic 48:252–262
Kerrigan Z, D'Hondt S (2022) Patterns of relative bacterial richness and community composition in seawater and marine sediment are robust for both operational taxonomic units and amplicon sequence variants. Front Microbiol 13:796758
Kishore D, Birzu G, Hu Z, DeLisi C, Korolev KS, Segrè D (2023) Inferring microbial co-occurrence networks from amplicon data: a systematic evaluation. mSystems 8(4):e00961
Kuypers MMM, Marchant HK, Kartal B (2018) The microbial nitrogen-cycling network. Nat Rev Microbiol 16:263-276
Lavrinienko A, Jernfors T, Koskimäki JJ, Pirttilä AM, Watts PC (2021) Does intraspecific variation in rDNA copy number affect analysis of microbial communities? Trends Microbiol 29(1):19–27
Leininger S, Urich T, Schloter M, Schwark L, Qi J, Nicol G et al (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442:806–809
Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F et al (2015) Determinants of community structure in the global plankton interactome. Science 348(6237):1262073
Liu C, Li C, Jiang Y, Zeng RJ, Yao M, Li X (2023) A guide for comparing microbial co-occurrence networks. iMeta 2(1):e71
Logares R, Deutschmann IM, Junger PC, Giner CR, Krabberød AK, Schmidt TSB et al (2020) Disentangling the mechanisms shaping the surface ocean microbiota. Microbiome 8:55
Luo H, Tolar BB, Swan BK, Zhang CL, Stepanauskas R, Ann Moran M et al (2014) Single-cell genomics shedding light on marine Thaumarchaeota diversification. ISME J 8(3):732–736
Lurgi M, Thomas T, Wemheuer B, Webster NS, Montoya JM (2019) Modularity and predicted functions of the global sponge-microbiome network. Nat Commun 10:992
Ma B, Wang H, Dsouza M, Lou J, He Y, Dai Z et al (2016) Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. ISME J 10(8):1891–1901
Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27(21):2957–2963
Massana R, DeLong EF, Pedrós-Alió C (2000) A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. Appl Environ Microbiol 66(5):1777–1787
Mo Y, Peng F, Gao X, Xiao P, Logares R, Jeppesen E et al (2021) Low shifts in salinity determined assembly processes and network stability of

microeukaryotic plankton communities in a subtropical urban reservoir. Microbiome 9:128
Moissl-Eichinger C, Pausan M, Taffner J, Berg G, Bang C, Schmitz RA (2018) Archaea are interactive components of complex microbiomes. Trends Microbiol 26(1):70–85
Mußmann M, Pjevac P, Krueger K, Dyksma S (2017) Genomic repertoire of the Woeseiaceae/JTB255, cosmopolitan and abundant core members of microbial communities in marine sediments. ISME J 11(5):1276–1281
Needham DM, Fuhrman JA (2016) Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. Nat Microbiol 1:16005
Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45(2):167–256
Newman MEJ (2006) Modularity and community structure in networks. Proc Natl Acad Sci USA 103(23):8577–8582
Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, Givens CE et al (2010) Genome characteristics of a generalist marine bacterial lineage. ISME J 4(6):784–798
Olesen JM, Bascompte J, Dupont YL, Jordano P (2007) The modularity of pollination networks. Proc Natl Acad Sci USA 104(50):19891–19896
Parada AE, Fuhrman JA (2017) Marine archaeal dynamics and interactions with the microbial community over 5 years from surface to seafloor. ISME J 11(11):2510–2525
Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P-A, Hugenholtz P (2021) GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Res 50(D1):D785–D794
Qian G, Li L, Hu X, Yu X, Ye L (2017) Enhancement of the biodegradability of activated sludge by the electric-coagulation multistage A/O membrane bioreactor treating low C/N industrial wastewater. Int Biodeterior Biodegrad 125:1–12
Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P et al (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596
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(5):1144–1158
Rinke C, Rubino F, Messer LF, Youssef N, Parks DH, Chuvochina M et al (2019) A phylogenomic and ecological analysis of the globally abundant Marine Group II archaea (*Ca*. Poseidoniales ord. nov.). ISME J 13(3):663–675
Spang A, Saw JH, Jorgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE et al (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature 521(7551):173–179
Takemoto K, Borjigin S (2011) Metabolic network modularity in Archaea depends on growth conditions. PLoS ONE 6(10):e25874
Tara Ocean Foundation, Tara Oceans, European Molecular Biology Laboratory (EMBL), European Marine Biological Resource Centre - European Research Infrastructure Consortium (EMBRC-ERIC) (2022) Priorities for ocean microbiome research. Nat Microbiol 7:937–947
Tilman D, Lehman CL, Thomson KT (1997) Plant diversity and ecosystem productivity: theoretical considerations. Proc Natl Acad Sci USA 94(5):1857–1861
Trevathan-Tackett SM, Sherman CDH, Huggett MJ, Campbell AH, Laverock B, Hurtado-McCormick V et al (2019) A horizon scan of priorities for coastal marine microbiome research. Nat Ecol Evol 3:1509–1520
Tully BJ (2019) Metabolic diversity within the globally abundant Marine Group II Euryarchaea offers insight into ecological patterns. Nat Commun 10:271
Wang K, Ye X, Chen H, Zhao Q, Hu C, He J et al (2015) Bacterial biogeography in the coastal waters of northern Zhejiang, East China Sea is highly controlled by spatially structured environmental gradients. Environ Microbiol 17(10):3898–3913
Wang K, Hu H, Yan H, Hou D, Wang Y, Dong P et al (2019a) Archaeal biogeography and interactions with microbial community across complex subtropical coastal waters. Mol Ecol 28(12):3101–3118
Wang Z, Juarez DL, Pan J-F, Blinebry SK, Gronniger J, Clark JS et al (2019b) Microbial communities across nearshore to offshore coastal transects are primarily shaped by distance and temperature. Environ Microbiol 21(10):3862–3872
Wang K, Yan H, Peng X, Hu H, Zhang H, Hou D et al (2020) Community assembly of bacteria and archaea in coastal waters governed by contrasting mechanisms: a seasonal perspective. Mol Ecol 29(19):3762–3776
Weiss S, Van Treuren W, Lozupone C, Faust K, Friedman J, Deng Y et al (2016) Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. ISME J 10(7):1669–1681
Wuchter C, Abbas B, Coolen MJL, Herfort L, van Bleijswijk J, Timmers P et al (2006) Archaeal nitrification in the ocean. Proc Natl Acad Sci USA 103(33):12317–12322
Yoshizawa S, Kumagai Y, Kim H, Ogura Y, Hayashi T, Iwasaki W et al (2014) Functional characterization of flavobacteria rhodopsins reveals a unique class of light-driven chloride pump in bacteria. Proc Natl Acad Sci USA 111(18):6732–6737
Zhang H, Huang X, Huang L, Bao F, Xiong S, Wang K et al (2018) Microeukaryotic biogeography in the typical subtropical coastal waters with multiple environmental gradients. Sci Total Environ 635:618–628
Zhou Z, Pan J, Wang F, Gu J, Li M (2018) Bathyarchaeota: globally distributed metabolic generalists in anoxic environments. FEMS Microbiol Rev 42(5):639–655
Zinger L, Boetius A, Ramette A (2014) Bacterial taxa-area and distance-decay relationships in marine environments. Mol Ecol 23(4):954–964

Publisher's Note