



## OPEN Heterogeneous selection dominated the temporal variation of the planktonic prokaryotic community during different seasons in the coastal waters of Bohai Bay

Wei Zhao<sup>1,2,3</sup>, Liuyang Shi<sup>1,2,3</sup>, Xingbiao Wang<sup>1,2</sup>, Jingjing Wang<sup>1,2</sup>, Song Xu<sup>1,2</sup>, Lei Ma<sup>1,2</sup>, Wanyu Zhu<sup>1,2</sup>, Xiaoxia Zhang<sup>1,2</sup>, Yifan Han<sup>1,2</sup> & Zhiyong Huang<sup>1,2</sup>✉

To explore temporal and spatial effects on the planktonic prokaryotic community composition (PCC) in the coastal region of the Bohai Sea, surface water samples were collected from 12 to 28 regularly distributed sites in Bohai Bay across 3 months from different seasons to characterize the PCC using high-throughput sequencing of the 16S rRNA V4 region. Prokaryotic  $\alpha$ - and  $\beta$ -diversity showed significant temporal variation during the three sampling months. VPA analysis based on both weighted and unweighted UniFrac distances exhibited a shift of environmental and spatial effects on PCC variation with temporal variation. Quantification analysis of assembly processes on community turn over showed that “heterogeneous selection” dominated for PCC temporal variation, with basic abiotic parameters such as temperature, pH, ammonia nitrogen as the driving factors. Analysis of seasonal features showed that seasonal specific OTUs (ssOTUs) exhibited different seasonal attributions under the same phylum; meanwhile, the ssOTUs showed significant correlations with the driving environmental factors, which suggested that finer-level analysis was needed to more strictly reflect the temporal variation. Moreover, predicted nitrogen and sulfur metabolism were significantly shifted during the temporal variation. Our results clearly showed that seasonally varied environmental factors drive the “heterogeneous selection” process for PCC assembly in seawaters of Bohai Bay during different sampling seasons.

With the development and utilization of diverse molecular ecology techniques based on prokaryotic 16S rRNA gene sequences, it has become clear that marine microbial communities undergo temporal and spatial variations similar to most planktonic assemblages<sup>1</sup>. The taxonomic composition of microbial assemblages and their temporal and spatial dynamics in the oceans are likely to be of major importance in determining the role of microorganisms in marine biogeochemistry<sup>2</sup>. The biogeographic or temporally ‘predictable’ patterns in the marine environment have been studied in coastal regions around the world<sup>3–9</sup>. Seasonal patterns have been observed in the San Pedro Channel, Bermuda Atlantic Time Series station, Western English Channel time series station and Hawaii Ocean time series station<sup>6</sup>. These long time series studies mostly concentrated on a certain location, while the spatial distribution of the microbial community was often affected by distances from coasts and terrestrial inputs<sup>7,10</sup>. When considering both the spatial and temporal effects on the community variation, different results were observed at different sites. In the northern Adriatic Sea, high variability characterizes the spatial structure of bacterial assemblages, while scarce seasonality was found at all nine studied stations, and the assemblages were generally strongly affected by river inputs, especially in spring, when freshwater loads

<sup>1</sup>Tianjin Key Laboratory of Industrial Biological Systems and Process Engineering, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China. <sup>2</sup>National Technology Innovation Center of Synthetic Biology, Tianjin, China. <sup>3</sup>These authors contributed equally: Wei Zhao and Liuyang Shi. ✉email: huang\_zy@tib.cas.cn

were higher<sup>1</sup>. Similarly, in coastal waters of the South China Sea, five regularly distributed sites in three seasons showed that the community diversity was better explained by spatial than by temporal patterns<sup>8</sup>. However, studies in Ofunato Bay (Pacific Ocean area of Japan)<sup>11</sup> and Gosung Bay (South Sea, Korea)<sup>9</sup> both showed that seasonal changes in environmental variables contribute to the dynamic structure of the bacterial community. It seems that the spatial and temporal distribution of marine microbial communities varies in different hydrological environments. Therefore, it is a challenge to draw a conclusion on how spatial and temporal factors determine the structuring of the microbial community in unexplored areas. Indeed, temporal factors are often considered to infer the environmental factors that change with time, such as temperature, DO, salinity and inorganic/organic resources, which inevitably change with the seasons. Therefore, rather than to say study the seasonal and spatial effect on microbial communities, it is better to say that it was to study the effects of environmental factors and spatial factors on microbial community over temporal variation.

When attempting to identify the key factors determining community variation, we need to confirm whether community variation was mainly influenced by deterministic or stochastic factors. Deterministic and stochastic processes influence the assembly of species into communities<sup>12,13</sup>. Deterministic processes are associated with ecological selection, in which abiotic and biotic factors determine the presence/absence and relative abundance of different species<sup>14</sup>. Stochastic processes include probabilistic dispersal and random changes in the relative abundance of different species (ecological drift) that are not caused by differences in environmentally determined fitness<sup>15</sup>. Deterministic and stochastic processes have been widely discussed in relation to various ecological processes, but little is known about their roles in the assembly of prokaryotic communities during natural temporal variation.

In addition, several studies have focused on the relationship between microbial community composition and ecosystem functioning<sup>1,16,17</sup>. Microbial communities are composed of functional groups (C and N fixers, ammonia oxidizers, sulfur reducers, etc.), so that the presence or absence of given organisms affects the biogeochemical activities occurring in a given environment, mainly by impacting elemental cycling rates<sup>1</sup>. Enrichment of sulfate-reducing bacteria was observed, and genes relevant for dissimilatory sulfate reduction were predicted in the coastal areas of the East China Sea<sup>16</sup>. The dynamics of nitrogen metabolism during the bloom process were also predicted, with different bacterial families contributing to the abundance of genes related to nitrogen metabolism<sup>17</sup>. However, the abundance of functional genes during different seasons in the coastal environment has rarely been researched, even though it may provide important information for understanding biogeochemical cycling. Moreover, changes in metabolic potential during different seasons also need to be taken into consideration.

This study was focused on Bohai Bay, China, which is located in the western region of the Bohai Sea in northern China. Based on the important geographical position and economic status of Bohai Bay, it has gradually become one of the intensive marine areas for marine scientific research. The spatial distribution of the bacterioplankton community in this area has been studied<sup>18</sup>, and it showed distinct spatial variation in the local environmental factors and dominant OTUs. Here, we selected 12–28 sites located at different geographic locations with different hydrological environments, among which several sites were located near the mouths of rivers, several sites were close to the tourist attractions, and several sites were at different distances from the shore. The main questions addressed by this study were as follows: (1) Do planktonic prokaryotic assemblages and functional potential show temporal patterns across different geographic locations? (2) How do environmental or spatial factors affect the PCC variation? (3) What processes drive the temporal or spatial pattern of the PCCs in the surface water of Bohai Bay?

## Results

**Variation in environmental parameters across space and time in Bohai Bay.** The environmental parameters of samples collected near the Tianjin coastal area from different stations and seasons exhibited high temporal and spatial heterogeneity. The seawater temperature was  $28.09 \pm 0.53$  °C in Aug,  $17.48 \pm 2.36$  °C in May, and  $19.55 \pm 1.26$  °C in Oct (Table 1). The seasonal variation in seawater temperature corresponded to the meteorological characteristics in Bohai Bay, with warm seawater in summer and relatively cool seawater in spring. The salinity was  $29.69 \pm 2.71$ ‰ in Aug,  $33.19 \pm 0.33$ ‰ in May, and  $30.15 \pm 1.63$ ‰ in Oct. Seasonal variations in salinity may be mainly related to freshwater loading. According to the precipitation observed data of Bohai Bay in previous years, the rainfall amount and days in summer are the most<sup>19</sup>, which may lead to the increase in runoff and the relatively low salinity in summer. Chlorophyll a (Chl a) was highest in May, with lower levels in Aug and Oct. The dissolved inorganic nitrogen (DIN) was significantly higher in May and Aug than in Oct. The higher level of DIN in May and Aug may be related to terrestrial input and supply for the demand of phytoplankton growth. In October, the temperature and DIN content were both not suitable for phytoplankton growth, and *Chl\_a* showed the lowest value. Spatially, the DIN distribution across the three seasons was rather similar, with high values observed in nearshore waters and low values in offshore waters (Dataset S1 & Fig. S1), which suggested that terrestrial input was an important source of DIN. The pH, soluble reactive phosphate (SRP) and chemical oxygen demand (COD) showed relatively higher values in October than in August and May, which may be caused by the dead phytoplankton release and terrestrial loadings through coasts and rivers. The dissolved oxygen (DO), conductivity and depth did not show significant variation among sampling times (Table 1), while the conductivity and depth had relatively higher values at offshore stations (Dataset S1) since the more remote the sampling water was, the greater the depth was in Bohai Bay and the closer it was to the open sea with higher salinity and conductivity. The ordination plot showed distinct partitioning of samples from nearshore and offshore sites along principal component axis 1 (PC1) (Fig. 1). The ordination plot could explain 73.49% of the total variation in the geo-physical-chemical parameters and revealed a linear positive correlation between different parameters (Fig. 1). AN, DIN, nitrate and *Chl\_a* were most crucial in the partitioning of samples from

	May	Aug	Oct
Temp/°C	17.48 ± 2.36 <sup>c</sup>	28.09 ± 0.53 <sup>a</sup>	19.55 ± 1.26 <sup>b</sup>
Salinity/‰	33.19 ± 0.33 <sup>a</sup>	29.69 ± 2.71 <sup>b</sup>	30.15 ± 1.63 <sup>b</sup>
Conductivity/ms/cm	47.04 ± 0.92 <sup>a</sup>	45.92 ± 2.44 <sup>a</sup>	46.49 ± 1.13 <sup>a</sup>
DO/mg/L	8.21 ± 0.51 <sup>a</sup>	7.58 ± 1.51 <sup>a</sup>	7.93 ± 0.39 <sup>a</sup>
PH	8.04 ± 0.07 <sup>b</sup>	8.39 ± 0.09 <sup>a</sup>	8.38 ± 0.07 <sup>a</sup>
SRP/mg/L	0.003 ± 0.002 <sup>c</sup>	0.01 ± 0.004 <sup>b</sup>	0.02 ± 0.01 <sup>a</sup>
COD/mg/L	0.3 ± 0.14 <sup>c</sup>	1.23 ± 0.82 <sup>b</sup>	2.45 ± 0.49 <sup>a</sup>
Nitrite/mg/L	0.02 ± 0.01 <sup>b</sup>	0.05 ± 0.04 <sup>a</sup>	0.04 ± 0.02 <sup>a</sup>
Nitrate/mg/L	0.24 ± 0.08 <sup>a</sup>	0.21 ± 0.09 <sup>a</sup>	0.16 ± 0.07 <sup>b</sup>
AN/mg/L	0.13 ± 0.02 <sup>a</sup>	0.1 ± 0.02 <sup>b</sup>	0.02 ± 0.02 <sup>c</sup>
DIN/mg/L	0.39 ± 0.08 <sup>a</sup>	0.36 ± 0.1 <sup>a</sup>	0.23 ± 0.07 <sup>b</sup>
un_ionN/mg/L	0.004 ± 0.001 <sup>b</sup>	0.01 ± 0.003 <sup>a</sup>	0.001 ± 0.001 <sup>c</sup>
Depth/m	8.31 ± 6.25 <sup>a</sup>	7.71 ± 6.01 <sup>a</sup>	6.98 ± 5.39 <sup>a</sup>
Trans/cm	0.79 ± 0.5 <sup>b</sup>	1.4 ± 0.65 <sup>a</sup>	1.2 ± 0.58 <sup>a</sup>
Chl_a/μg/L	4.84 ± 2.76 <sup>a</sup>	2.37 ± 1.77 <sup>b</sup>	2.24 ± 1.74 <sup>b</sup>
Faith_pd	58.28 ± 12.75 <sup>b</sup>	67.99 ± 16.77 <sup>a</sup>	44.36 ± 3.68 <sup>c</sup>
OTU_richness	863.75 ± 211.61 <sup>b</sup>	985.37 ± 256.73 <sup>a</sup>	675.01 ± 58.14 <sup>c</sup>
Shannon	7.37 ± 0.44 <sup>c</sup>	7.91 ± 0.69 <sup>a</sup>	7.67 ± 0.28 <sup>b</sup>
Evenness	0.76 ± 0.03 <sup>c</sup>	0.8 ± 0.04 <sup>b</sup>	0.82 ± 0.02 <sup>a</sup>

**Table 1.** The independent-samples *t* test of environmental variables and  $\alpha$ -diversity among different months. Data in the table are the Means  $\pm$  SD; The different normal letters (abc) in the same row indicate significant difference among sampling times at 0.05 level ( $n > 3$ ), if they contain the same letter it means the difference is not significant, if they contain different letter it means the difference was significant. Temp, water temperature; DO, dissolved oxygen; SRP, soluble reactive phosphates; COD, chemical oxygen demand; AN: ammonia nitrogen; DIN, dissolved inorganic nitrogen; Trans, transparency; un\_ionN, nonionic ammonia; Chl\_a, chlorophyll a.

May and the other 2 months; salinity, longitude, depth and conductivity were crucial for the partitioning of samples from offshore and nearshore stations; pH, COD, SRP, nitrite and temperature were crucial for the partitioning of samples from nearshore stations in August and October and samples from offshore stations. Overall, the principal component analysis (PCA) plot clearly showed both the temporal and spatial variation of the measured environmental parameters, indicating that complex biogeochemical processes and hydrodynamic conditions lead to the variation among sites and seasons.

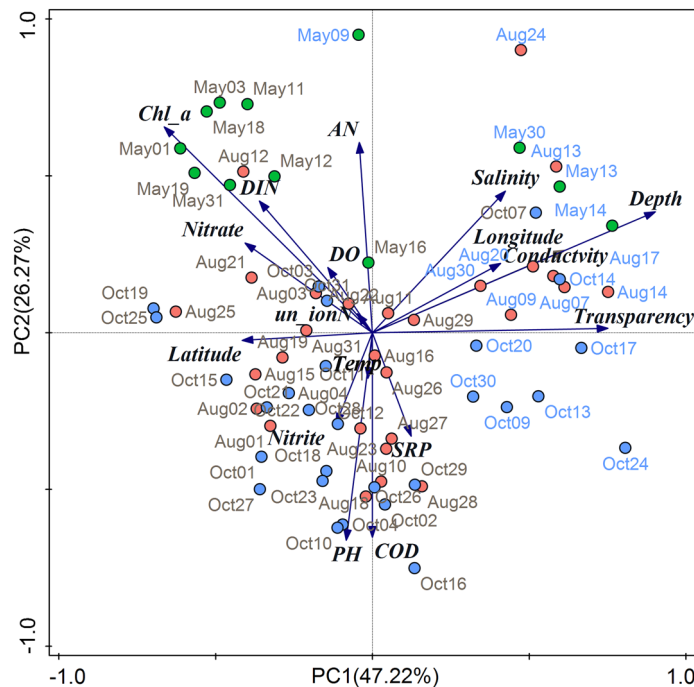
**Prokaryotic  $\alpha/\beta$ -diversity variation.** Measures of  $\alpha$ -diversity showed significant differences in shannon, evenness, faith\_pd and OTU richness between samples from May/Aug and Oct (Fig. 2, Table 1). Principal coordinates analysis (PCoAs) based on weighted UniFrac (WUF) distance and unweighted UniFrac (UUF) distance showed that the PCC from different sampling months separated across the first and second principal coordinates (Fig. 3A-B). Both the analysis of similarity (ANOSIM) and permutational multivariate analysis of variance (PERMANOVA/ADONIS) results indicated that the prokaryotic communities varied significantly across different sampling months when using a WUF distance metric (ANOSIM,  $r = 0.709$ ,  $P = 0.001$ ; ADONIS,  $R^2 = 40.0\%$ ,  $P = 0.001$ ) and UUF distance metric (ANOSIM,  $r = 0.934$ ,  $P = 0.001$ ; ADONIS,  $R^2 = 38.7\%$ ,  $P = 0.001$ ). At the same time, the prokaryotic  $\alpha$ - and  $\beta$ -diversity both showed high within-month variability in Aug (Figs. 2, 3C-D), which indicated that the community varied greatly among different sites in Aug.

**Correlation between prokaryotic  $\alpha/\beta$ -diversity and physical, chemical and geographic factors.** The  $\alpha$ -diversity measurements exhibited significant positive correlations with temperature, pH, SRP, AN and un\_ionN (Dataset S2). The correlation between  $\alpha$ -diversity indexes and geo factors (longitude and latitude) was not strong or significant both in samples across the three sampling times or from each sampling time (Dataset S2).

The environmental variation significantly correlated with  $\beta$ -diversity among the three seasons ( $r_{\text{weighted}} = 0.4558$ ,  $r_{\text{unweighted}} = 0.4631$ ,  $P = 0.001$ , Table 2), with pH, AN, temperature, un\_ionN, COD, nitrite, SRP, salinity, DO and DIN as the main individual determinants. However, it did not show significant correlations with  $\beta$ -diversity at any sampling time except in Oct (Table S1).

The geographic distance was not correlated with prokaryotic  $\beta$ -diversity (variation in community composition;  $r < 0.03$ ,  $P > 0.05$ ; Table 2) among the three sampling times. However, samples from Aug and Oct exhibited a significant correlation between  $\beta$ -diversity and geographic distance (Table S1).

**Factors driving the PCC variation.** PERMANOVA using the UUF/WUF distance indicated that temperature variation explained the largest part of community variation among the investigated factors (34.90%/19.83%,



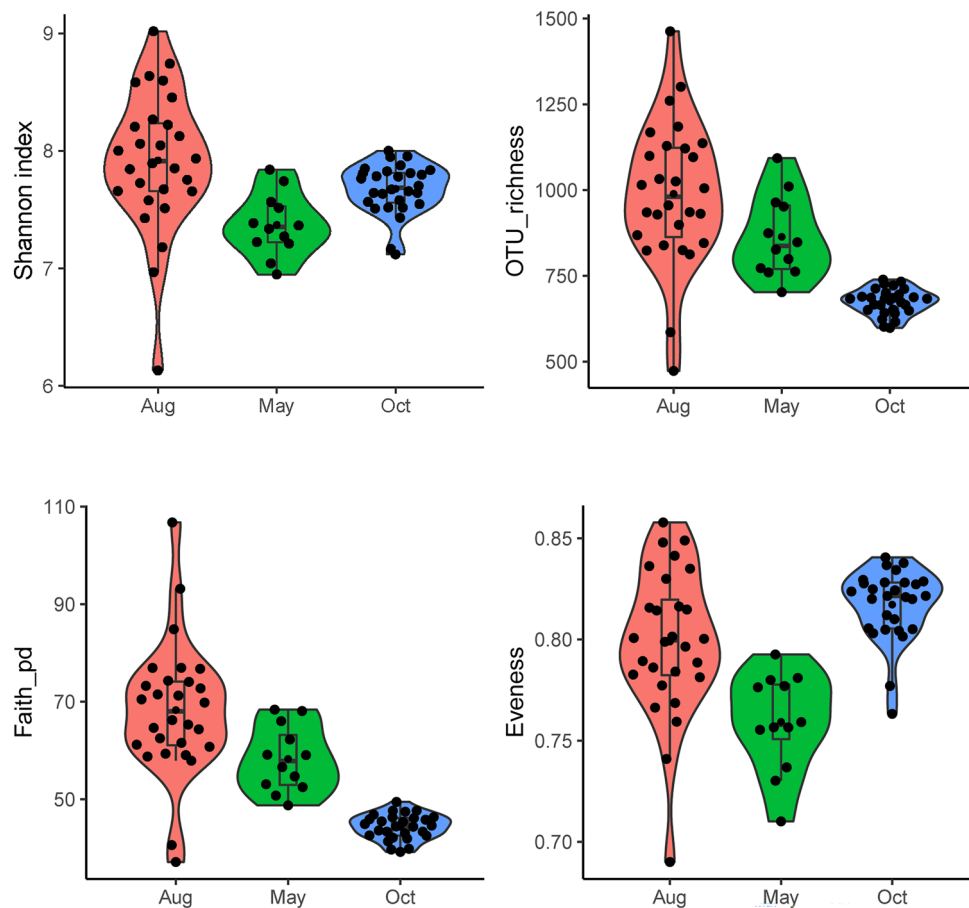
**Figure 1.** Biplot of the principal component analysis (PCA) for environmental parameters in the seawater samples of the Bohai Bay coastal area across different seasons and sites. The two principal components (PC1 and PC2) explained 73.49% of the total variation in the environmental data and showed clear partitioning of offshore samples (in blue font) from other nearshore samples along PC1 and partitioning of May samples from August and October along PC2. The variables transparency and latitude were strongly correlated with PC1, and the variables ammonia nitrogen (AN), COD, pH, soluble reactive phosphate (SRP), and nitrite were strongly correlated with PC2. Chlorophyll a (*Chl\_a*), dissolved inorganic nitrogen (DIN), nitrate and DO were mainly positively correlated with samples from May, while salinity, longitude, depth and conductivity were mainly positively correlated with offshore samples. Blue arrows represent environmental parameters, and circles in color represent sampling points.

$P = 0.001$ , Dataset S3), with AN (31.84%/13.56%,  $P = 0.001$ ) and salinity (12.91%/6.21%,  $P = 0.001$ ) as the second and third most significant sources of variation.

The variance partitioning analysis (VPA) conducted on both UUF/WUF distances showed that almost 100% percent of the variation in PCC among all three sampling times was explained by the detected environmental factors. In May, no environmental or spatial factors could be selected as significantly explain the PCC variation; in Aug, the joint effects of environmental and spatial factors could explain 49.5% of the variation; in Oct, based on WUF distance, the spatial factors could purely explain 10.5%, environmental factors could purely explain 38.8%, their joint effects could explain 28.2%, and based on UUF distance, the joint effects of environmental factors and trend could explain 13.7% of the PCC variation. These results indicated dramatic shifts in the spatial or environmental factor effects on the PCC variation at different sampling times in Bohai Bay (Table 3).

**Distinct seasonal features at the phylum and OTU levels.** There were notable differences in the proportions of various phyla among different seasons (sampling month). In May, there was a greater proportion of Alphaproteobacteria (41.41%), Planctomycetes (6.42%), Actinobacteria (3.86%), Firmicutes (1.48%), Acidobacteria (0.45%), TM7 (0.16%), Tenericutes (0.16%), OD1 (0.13%), and WPS-2 (0.09%) than in Aug and Oct, whereas Gammaproteobacteria (44.23%), GN02 (0.08%) and SAR406 (0.04%) were depleted in May and Aug but enriched in Oct. In Aug, Bacteroidetes (13.98%), Deltaproteobacteria (6.93%), Verrucomicrobia (4.5%), Chloroflexi (0.36%), Lentisphaerae (0.97%), TM6 (0.25%), Nitrospirae (0.08%), Chlamydiae (0.07%), Chlorobi (0.07%), Spirochaetes (0.04%) and OP8 (0.03%) were significantly enriched than in the other two sampling times (Duncan test; Table S2).

At the OTU level, OTUs with relative abundance  $> 0.01\%$  (1040 OTUs) were used to perform the difference analysis, and 175 OTUs in May, 281 OTUs in Aug, and 210 OTUs in Oct were identified as seasonal specific OTUs (ssOTUs). The cooccurrence network showed that the ssOTUs were clustered separately (Fig. 4A). Furthermore, the separation of the three modules contained most of the ssOTUs specific to different seasons (Fig. 4A–B). All the ssOTUs of different seasons comprised a taxonomically broad set of prokaryotes at the phylum (phylum Proteobacteria is grouped at the class level) level (Fig. 4C) belonging to various phyla with different proportions. Betaproteobacteria, Verrucomicrobia, Gemmatimonadetes, Epsilonproteobacteria, PAUC34f., and Euryarchaeota did not show significant differences among the three sampling times at the phylum level, but features belonging to

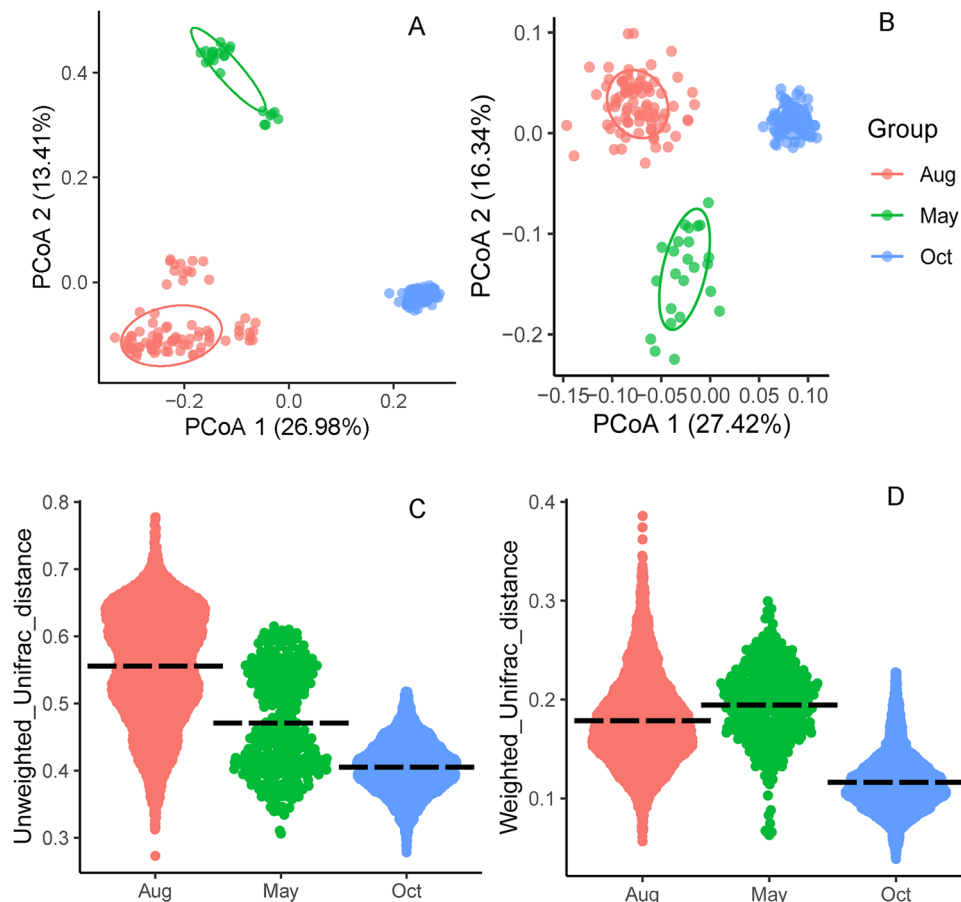


**Figure 2.** Alpha diversity of shannon, evenness, faith\_pd (phylogenetic diversity) and OTU richness value of the prokaryotic community of all the samples from different stations at different sampling times.

these phyla showed differences at the OTU level (Fig. 4C, Dataset S4). In addition, the phylum ssOTUs belonging to, such as Alphaproteobacteria, Gammaproteobacteria, Bacteroidetes, Actinobacteria, and Deltaproteobacteria, were not only enriched at one sampling time (Dataset S4) but also enriched at the other two sampling times (Fig. 4C, Dataset S4). These results revealed that different seasons do not strictly select specific microbial lineages at the phylum level, but a finer level analysis could more strictly reflect the seasonal variation.

Regression analysis between the relative abundance of modules to which the ssOTUs belonged and the environmental factors was also conducted, and module 1 abundance, to which the Aug-ssOTUs belonged, showed a significant positive correlation with temperature ( $R^2 = 0.77$ ,  $P = 6.609 \times 10^{-62}$ ), AN ( $R^2 = 0.43$ ,  $P = 7.416 \times 10^{-25}$ ), and un\_ionN ( $R^2 = 0.75$ ,  $P = 1.366 \times 10^{-58}$ ) and a negative correlation with SRP ( $R^2 = 0.81$ ,  $P = 6.762 \times 10^{-17}$ ). This may be caused by the functional role of the microbes in Aug. In the Aug-ssOTUs, Deltaproteobacteria showed a higher ratio than in the other 2 months (Fig. 4c), and in the following functional analysis, Deltaproteobacteria contributed to the genes related to nitrogen fixation, which may help to explain why there was a positive correlation of Aug-ssOTUs to AN and un\_ionN. The module 2 abundance to which the May-ssOTUs belonged showed a significant negative correlation with pH ( $R^2 = 0.65$ ,  $P = 4.026 \times 10^{-44}$ ), temperature ( $R^2 = 0.19$ ,  $P = 2.325 \times 10^{-10}$ ), un\_ionN ( $R^2 = 0.025$ ,  $P = 0.01779$ ), and SRP ( $R^2 = 0.12$ ,  $P = 4.104 \times 10^{-07}$ ) and a positive correlation with AN ( $R^2 = 0.26$ ,  $P = 5.174 \times 10^{-14}$ ). In the May-ssOTUs, the ratio of Alphaproteobacteria was the highest, and Alphaproteobacteria were reported to be pH-sensitive groups in marine environments<sup>20</sup>, which prefer neutral pH environments<sup>21</sup>. In this study, the pH in May was  $8.04 \pm 0.07$ , in Aug was  $8.39 \pm 0.09$ , in Oct was  $8.38 \pm 0.07$ , and the pH in May was the closest to neutral, and the ratio decreased with increasing pH in Oct and Aug. The abundance of module 3, to which the Oct-ssOTUs belonged, showed a significant positive correlation with SRP ( $R^2 = 0.81$ ,  $P = 0.16 \times 10^{-10}$ ) and pH ( $R^2 = 0.054$ ,  $P = 0.00075$ ) and a negative correlation with temperature ( $R^2 = 0.44$ ,  $P = 2.276 \times 10^{-25}$ ), AN ( $R^2 = 0.75$ ,  $P = 4.51 \times 10^{-58}$ ), and un\_ionN ( $R^2 = 0.6$ ,  $P = 3.995 \times 10^{-39}$ ) (Fig. S2). Phosphate has been identified to limit primary productivity<sup>22</sup>, which is of great importance in the structure of dominant bacterial taxa in marine environments<sup>23</sup>. In the Oct-ssOTUs, the ratio of Gammaproteobacteria was the highest, as reported. Gammaproteobacteria was significantly explained by SRP during the seasonal variation in the Western English Channel, with Rho equal to 0.75<sup>23</sup>, which suggested the sensitivity of it to SRP, and in that study, it also showed a negative correlation between temperature and Gammaproteobacteria and a positive correlation between SRP and Gammaproteobacteria. Although the correlation was not significant, the variation trend was consistent, which indicates that the phenomenon observed in this study was not unexpected. In addition, most ammonia-oxidizing bacteria belong





**Figure 3.** Principal coordinate analysis (PCoA) based on unweighted (A) and weighted (B) UniFrac distances for prokaryotic communities in the surface waters; box plots showing the unweighted (C) and weighted (D) UniFrac distances among each station at different sampling times.

to the Betaproteobacteria and Gammaproteobacteria classes are chemolithoautotrophs that oxidize ammonia to nitrite<sup>24</sup>. Gammaproteobacteria and Betaproteobacteria both had higher ratios in Oct-ssOTUs, and the functional prediction results also showed that *pmoA/amoA* and *pmoB/amoB*, which encode ammonia monooxygenase, were mainly contributed by OTUs from Gammaproteobacteria and Betaproteobacteria (Dataset S10). The utilization of ammonia may explain the negative correlation between the Oct-ssOTUs and AN.

**Community assembly processes across different sampling months and sites.** Based on the analysis of phylogenetic turnover, unweighted  $\beta$ NTI mostly ranged from -2 to 2 across different sites at a single sampling time in May, Aug and Oct, revealing that PCC variations across different sampling sites at a single time were mostly affected by stochastic processes. The unweighted  $\beta$ NTI was greater than 2 across May–Aug, May–Oct and Aug–Oct (Fig. 5A), which revealed that the variations in PCC across different sampling times were mostly affected by deterministic processes. The RCbray values across any two sampling times were equal to 1, and in each sampling time, the RCbray values ranged from -1 to 1 (Fig. 5B). Combining the  $\beta$ NTI and RCbray values, the community assembly processes were quantified at each sampling time and at any two sampling times. As shown in Fig. 5C, turning over of the community during different sampling times was mainly governed by selection; among the different sites in May and Oct, it was mainly governed by “undominated” processes; community turn over in Aug was mainly governed by the influence of “Dispersal Limitation”. These results indicated that the shifts in the assembly of prokaryotic communities during different sampling times were caused by strong “heterogeneous selection” ( $\beta$ NTI > 2), and the community variation at each sampling time was mainly caused by stochastic processes.

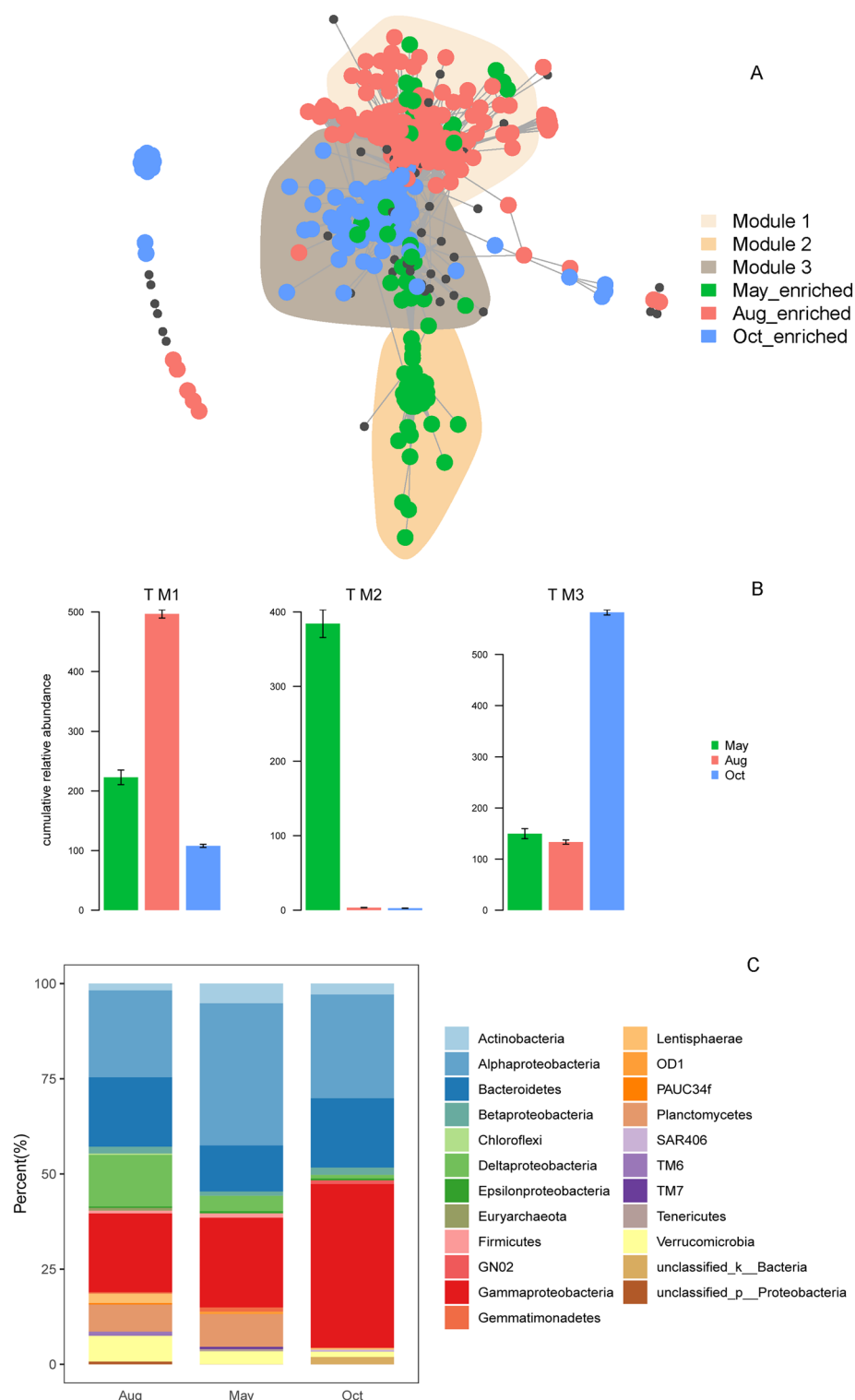
**Prediction of the metabolic potential at different sampling times.** The NSTI scores of each sample ranged from 0.033 to 0.096, with a mean of 0.058 (Dataset S5). Microbial functions were detected in all the samples from the three sampling times, and it was found that the relative abundances of 242 pathways were significantly changed between samples from May and samples from Aug (Dataset S6). The relative abundances of 321 pathways were significantly changed between samples from May and Oct (Dataset S7), and the relative abundances of 370 pathways were significantly changed between samples from Aug and Oct (Dataset S8).

Variables	Weighted					Unweighted				
	r	p	r <sub>-temporal</sub>	p <sub>-temporal</sub>	N	r	p	r <sub>-temporal</sub>	p <sub>-temporal</sub>	N
pH	<b>0.4392</b>	0.001	0.05697	0.144	68	<b>0.5104</b>	0.001	− 0.1392	1	68
AN	<b>0.4939</b>	0.001	<b>0.1403</b>	0.004	68	<b>0.5093</b>	0.001	<b>0.1922</b>	0.001	68
Env	<b>0.4558</b>	0.001	<b>0.1964</b>	0.002	68	<b>0.4631</b>	0.001	<b>0.2551</b>	0.001	68
Temp	<b>0.4819</b>	0.001	− 0.0596	0.854	68	<b>0.5491</b>	0.001	− 0.004416	0.554	68
un_ionN	<b>0.3849</b>	0.001	0.0199	0.403	68	<b>0.4657</b>	0.001	<b>0.2675</b>	0.001	68
COD	<b>0.3587</b>	0.001	<b>0.1043</b>	0.007	68	<b>0.3903</b>	0.001	<b>0.1848</b>	0.001	68
Nitrite	<b>0.2518</b>	0.001	0.155	0.098	68	<b>0.1887</b>	0.002	<b>0.1229</b>	0.01	68
SRP	<b>0.133</b>	0.009	− 0.05007	0.803	68	<b>0.1911</b>	0.002	<b>0.09605</b>	0.026	68
Salinity	<b>0.09235</b>	0.046	− 0.02944	0.653	68	<b>0.1855</b>	0.001	<b>0.0967</b>	0.02	68
DO	<b>0.1653</b>	0.014	0.06852	0.201	68	<b>0.2042</b>	0.001	<b>0.2661</b>	0.001	68
DIN	<b>0.2157</b>	0.002	<b>0.1239</b>	0.031	68	<b>0.173</b>	0.001	<b>0.09299</b>	0.035	68
Chl <sub>a</sub>	0.09864	0.076	− 0.03354	0.688	68	<b>0.1867</b>	0.001	− 0.04271	0.839	68
Nitrate	0.1007	0.061	0.092	0.09	68	0.05499	0.114	0.01883	0.289	68
Depth	0.0925	0.071	0.1442	0.018	68	− 0.04877	0.823	0.002255	0.421	68
Trans	0.1017	0.019	<b>0.1477</b>	0.022	68	0.01094	0.35	− 0.02779	0.732	68
Geo	0.03149	0.293	0.03907	0.295	68	− 0.005656	0.491	0.03601	0.233	68
Longitude	0.02462	0.331	0.0302	0.327	68	− 0.02081	0.644	0.013	0.379	68
Latitude	− 0.0125	0.603	0.01818	0.326	68	0.0194	0.25	<b>0.08091</b>	0.014	68
Conductivity	− 0.00484	0.503	0.06671	0.176	68	− 0.00338	0.473	<b>0.1746</b>	0.002	68

**Table 2.** Spearman's rank correlation between environmental/spatial variability (Euclidean distance) and prokaryotic  $\beta$ -diversity (weighted/unweighted UniFrac distance) among all samples from different season. r, correlation coefficients between environmental/spatial variability and prokaryotic community distance derived from Mantel testes with 999 permutations,  $r_{\text{-temporal}}$ , correlation coefficients between environmental/spatial variability and prokaryotic community distance controlled by temporal variation derived from partial Mantel testes with 999 permutations. Data in bold indicate significant correlations ( $P < 0.05$ ). Abbreviations: Env, environmental distance (Euclidean distance); Geo, geographic distance. Refer to Table 1 for other abbreviations.

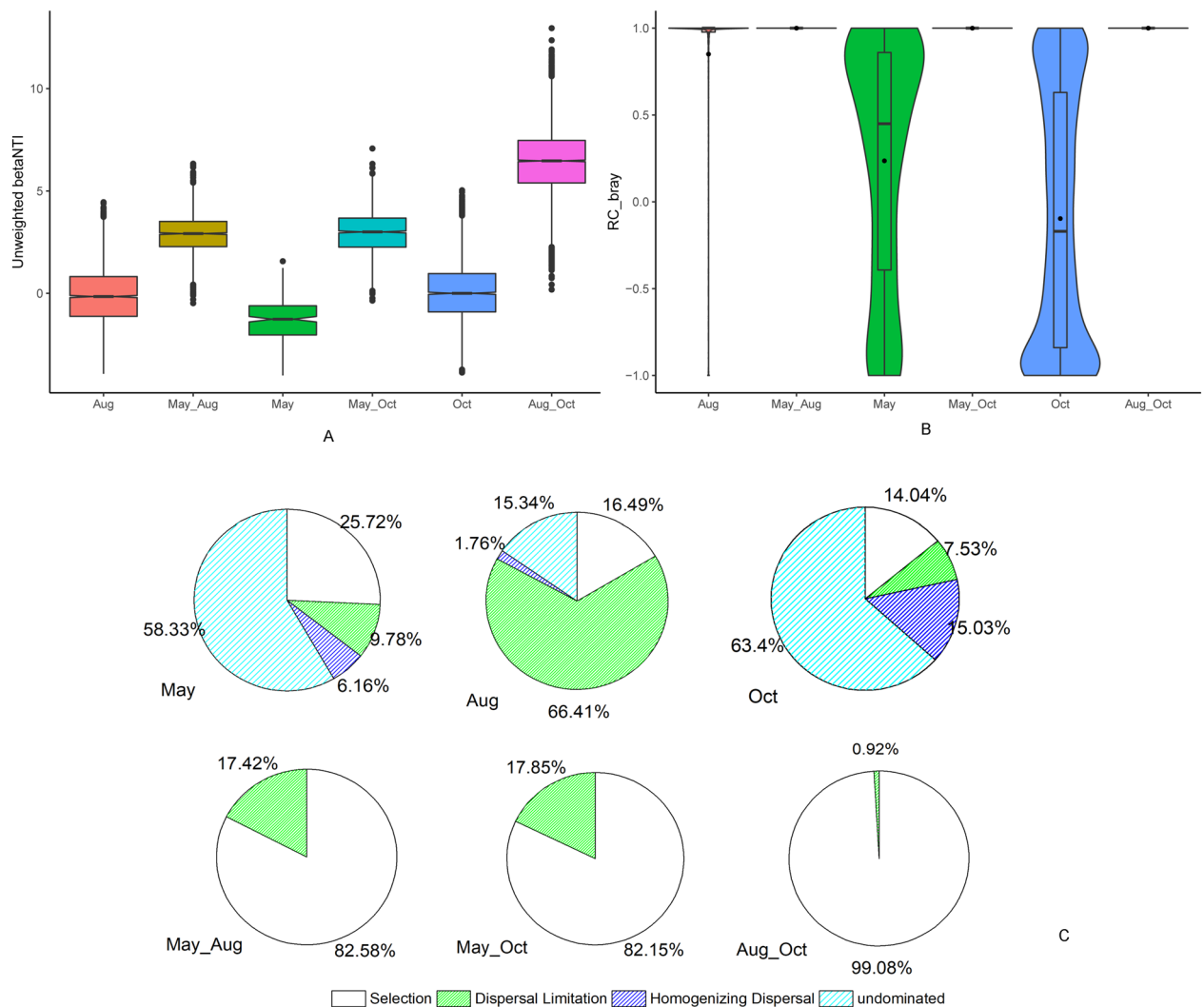
	Env	Trend	PCNM	Trend*Env	Env*PCNM	Trend*PCNM	Trend*Env*PCNM
<b>May</b>							
Weighted	✕	✕	✕	✕	✕	✕	✕
Unweighted	✕	✕	✕	✕	✕	✕	✕
<b>Aug</b>							
Weighted	0.114 (0.051)	0.019 (0.277)	0 (0.411)	0.082	0.067	0.078	0.095
Unweighted	0.082 (0.073)	0 (0.752)	0.053 (0.095)	0.032	0.189	0.112	0.162
<b>Oct</b>							
Weighted	0.388 (0.001)	0.028 (0.051)	0.105 (0.002)	0	0.016	0.021	0.245
Unweighted	0.077 (0.051)	0.06 (0.038)	0.007 (0.286)	0.137	0.08	0.033	0.037
<b>All</b>							
Weighted	0.995 (0.001)	0 (0.453)	0 (0.749)	0	0	0	0
Unweighted	1.042* (0.001)	0.001 (0.28)	0 (0.861)	0	0	0.001	0

**Table 3.** Variance partitioning analysis of prokaryotic community in Bohai Bay according to seawater environmental factors and geospatial factors. The spatial factors including linear trend and PCNM variables. Forward selection procedures were used to select the best subset of environmental, trend, and PCNM variables explaining community variation, respectively. The community variation was calculated on the weighted and unweighted UniFrac distance matrix, respectively. Monte Carlo permutation test was performed on each set without the effect of the other by permuting samples freely (999 permutations). Values in the table represent how many percent could the variables explain the PCC variation and values in parentheses represent the  $P$  value based on Monte Carlo permutation test. The asterisk (\*) which the individual fractions of Env accounted more than 1 since there were negative  $R^2_{\text{adj}}$  value during the redundancy analysis, negative  $R^2_{\text{adj}}$  means that the explanatory variable can account for a smaller proportion of the total variation than that of the randomly generated variable. Therefore, usually in practical application, negative  $R^2_{\text{adj}}$  is usually treated as 0, but it must be labeled as true  $R^2_{\text{adj}}$  in order for all  $R^2_{\text{adj}}$  components to add up to 1 (Referenced from Legendre and Legendre, 2012).



**Figure 4.** Co-occurrence patterns of seasonal sensitive OTUs (A). Co-occurrence network visualizing significant correlations ( $p > 0.7$ ,  $P < 0.001$ ; indicated with grey lines) between OTUs from the three sampling times with abundance  $> 0.01\%$ . Different colors represent ssOTUs in May (green), Aug (red) and Oct (blue). Cumulative relative abundance (as counts per million, CPM; y-axis in  $\times 1000$ ) of all the sensitive modules in the networks (B). The phylum attribution of ssOTUs in each season (C). The y-axis is the percentage of the number of OTUs that belong to a particular phylum that accounts for the total number of all the OTUs.





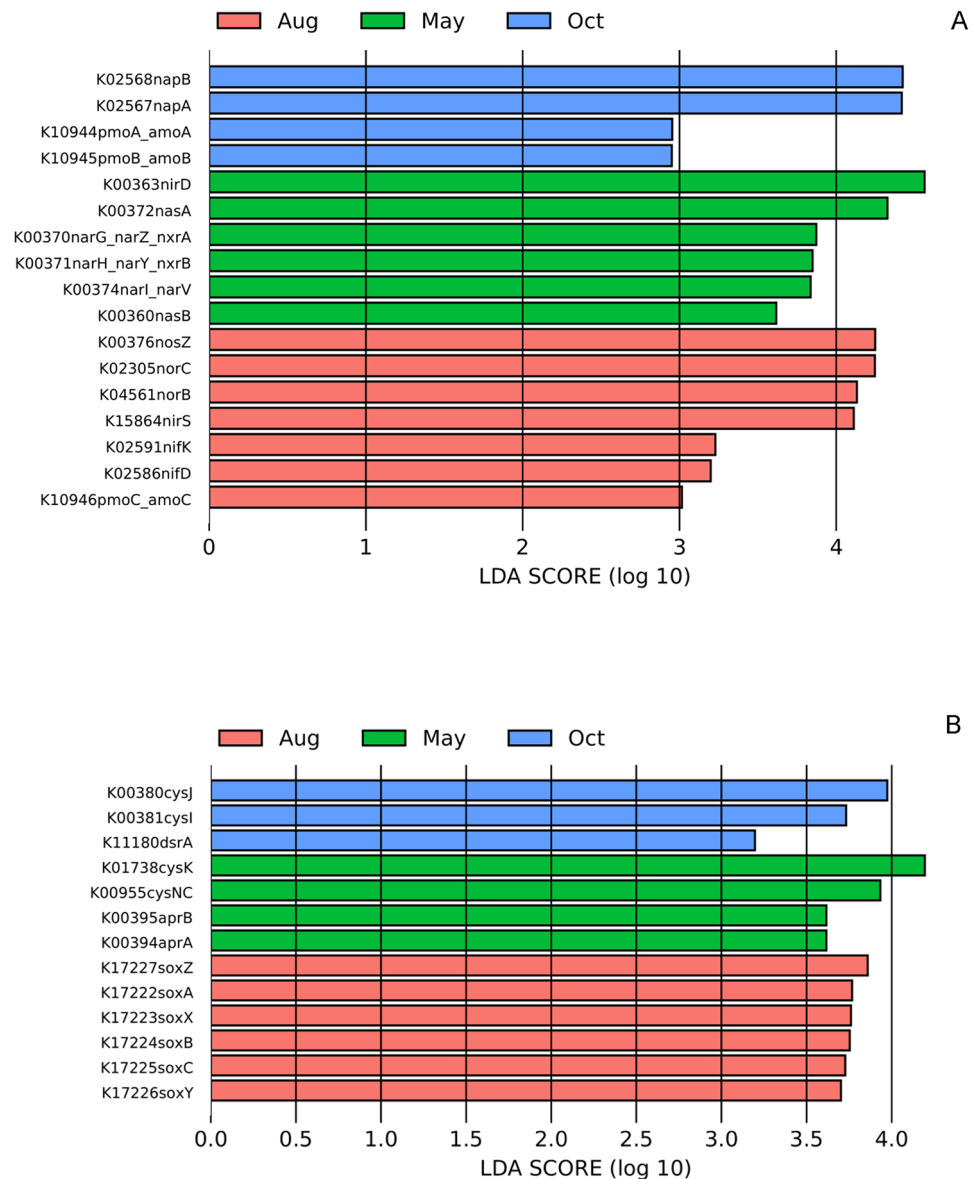
**Figure 5.** Patterns of distribution of unweighted  $\beta$ NTI (A) and RCbray (B) values across different sampling times. Quantification of the features that impose community assembly processes in and among different sampling times. (C) Pie charts give the percent of turnover in community composition governed primarily by Selection acting alone (white fill), Dispersal Limitation (green line fill), Homogenizing Dispersal (blue line fill) and undominated process (cyan fill).

Genes related to energy metabolism were given more attention. For nitrogen metabolism genes relevant with nitrogen fixation (*nifD*, *nifK*) were detected only enriched in Aug, while genes relevant with nitrate reduction and denitrification (*narG*, *narZ*, *nxrA*, *narH*, *narY*, *nxB*, *narL*, *narV*, *nirD*, *nasA*, *nasB*) were detected enriched in May, genes related with ammonia oxidation were both detected enriched in Oct and Aug. For sulfur metabolism, genes relevant with thiosulfate oxidation (*soxA*, *soxB*, *soxC*, *soxX*, *soxY* and *soxZ*) were only enriched in Aug, while genes relevant with sulfate and sulfite reduction (*cysNC*, *aprA*, *aprB*, *cysJ*, *cysI*, *cysK*, *dsrA*) were detected enriched in May and Oct (Fig. 6).

Prokaryotic taxa contributed to the significantly varied functional genes related to nitrogen and sulfur metabolism at different sampling times. At the species level, the taxa contributing to *nifK* and *nifD* mainly belonged to Deltaproteobacteria and Firmicutes, and the taxa contributing to the *sox*-series genes mainly belonged to Alphaproteobacteria and Gammaproteobacteria (Fig. S3). The denitrification-related functional genes that were enriched in May were mainly contributed by members from Alphaproteobacteria and Gammaproteobacteria. The taxa contributing to *dsrA*, *aprA* and *aprB* were mainly from Deltaproteobacteria, including members of Desulfarculaceae, Desulfobacteraceae, Desulfobulbaceae, Desulfobacteraceae and Syntrophobacteraceae (Fig. S4).

## Discussion

We found clear temporal variation in the composition of the prokaryotic community in the surface water of Bohai Bay across different sites during the three sampling months. We observed that environmental factors have a stronger effect on PCC variation than spatial factors during different sampling times. Overall, environmental factors such as temperature, pH, AN, un\_ionN and SRP all showed significant correlations with  $\alpha/\beta$ -diversity (Dataset S2, Table 2). In addition, the coefficients between environmental factors (temperature, pH, AN, un\_ionN, COD, SRP, DO, DIN) and



**Figure 6.** The LefSe analysis indicated significantly differential abundances of PICRUSt predicted genes relevant to energy metabolism in different months of samples.

$\beta$ -diversity/ $\beta$ NTI were decreased when the temporal factor was deconvoluted (Tables 2, 4), which suggested that these factors were also seasonally varied, and it can be speculated that seasonally varied environmental factors contributed to the dramatic prokaryotic temporal variation. Temperature plays a fundamental role in regulating the activity and growth of microorganisms, and changes in temperature could have a major direct effect on the population dynamics of marine bacteria<sup>25</sup>. In this study, temperature had the highest correlation coefficient with unweighted  $\beta$ -diversity (Table 2), which indicated its important role in affecting seasonal community variation. pH is also an important factor in shaping marine microbial communities<sup>20,26,27</sup>. Many studies have focused on the influence of pH decrease as a consequence of ocean acidification on marine zooplankton<sup>26,28,29</sup>. Acidification may impact prokaryotic communities by shifting towards plankton dominance, affecting productivity and nutrient cycling as well as increasing the abundance of disease-causing bacteria<sup>20,28</sup>. In contrast to previous studies, the community shift with pH variation in this study seemed not to be caused by ocean acidification, since the previously reported increase in the abundance of Flavobacteriales (Bacteroidetes)<sup>20</sup> and disease-causing bacteria, such as Vibrionaceae and Alteromonadaceae (Gammaproteobacteria)<sup>26,28</sup>, at lower pH values was not observed in Bohai Bay in May (Table S2). Conversely, the pH observed in this study was slightly alkaline, with which in Aug was the highest. The rising temperature and strong photosynthesis in Aug may reduce the concentration of carbon dioxide and hydrogen ions in the surface seawater, resulting in a rise in pH. In addition, the riverine input and anthropogenic activities during different seasons may also affect the pH value. AN was reported to be an important factor in the marine environment affecting the variation in OTU richness<sup>30</sup>. In a previous study of the same research scope in Bohai Bay<sup>18</sup>, AN showed a significant correlation with the  $\beta$ -diversity of the bacterioplankton community along with differently polluted areas. In this study, there was a positive correlation between AN and OTU richness and

Variables	Weighted					Unweighted				
	r	p	r <sub>temporal</sub>	p <sub>temporal</sub>	N	r	p	r <sub>temporal</sub>	p <sub>temporal</sub>	N
pH	− 0.1049	0.975	− 0.3321	1	68	0.02956	0.142	− 0.6056	1	68
AN	<b>0.3439</b>	0.001	<b>0.1858</b>	0.001	68	<b>0.5913</b>	0.001	<b>0.2918</b>	0.001	68
Env	<b>0.1673</b>	0.007	0.04493	0.231	68	<b>0.3306</b>	0.001	0.03383	0.135	68
Temp	<b>0.3442</b>	0.001	<b>0.2275</b>	0.001	68	<b>0.6558</b>	0.001	<b>0.4364</b>	0.001	68
un_ionN	<b>0.4053</b>	0.001	<b>0.3115</b>	0.001	68	<b>0.7201</b>	0.001	<b>0.6174</b>	0.001	68
COD	<b>0.2293</b>	0.001	<b>0.1245</b>	0.002	68	<b>0.3853</b>	0.001	<b>0.1345</b>	0.002	68
Nitrite	0.08334	0.05	0.07562	0.125	68	0.003717	0.385	− 0.0406	0.888	68
SRP	<b>0.2698</b>	0.001	<b>0.1957</b>	0.002	68	<b>0.3919</b>	0.001	<b>0.321</b>	0.001	68
Salinity	− 0.02396	0.671	− 0.1029	0.988	68	0.02563	0.123	− 0.1433	1	68
DO	<b>0.102</b>	0.042	0.006755	0.441	68	<b>0.2327</b>	0.001	<b>0.123</b>	0.001	68
DIN	<b>0.09197</b>	0.038	0.01272	0.4	68	<b>0.1929</b>	0.001	<b>0.08925</b>	0.005	68
Chlorophyll_a	− 0.1087	0.985	− 0.1467	0.998	68	− 0.003634	0.54	− 0.202	1	68
Nitrate_N	− 0.003856	0.555	− 0.03669	0.716	68	0.009555	0.35	− 0.04097	0.926	68
Depth	0.03468	0.245	0.017	0.327	68	− 0.0894	1	− 0.1126	1	68
Trans	− 0.02311	0.708	− 0.03759	0.747	68	− 0.007616	0.608	− 0.08974	1	68
Geo	0.04925	0.151	0.04158	0.241	68	− 0.05309	0.98	− 0.07174	0.989	68
Longitude	0.03519	0.265	0.03333	0.284	68	− 0.04764	0.97	− 0.06673	0.981	68
Latitude	0.00531	0.397	0.02266	0.28	68	− 0.01947	0.819	− 0.01521	0.721	68
Conductivity	− 0.09627	0.969	− 0.1104	0.968	68	0.03493	0.105	0.02299	0.191	68

**Table 4.** Spearman's rank correlation between environmental/spatial variability (Euclidean distance) and weighted and unweighted  $\beta$ NTI among all samples from different season. r, correlation coefficients between environmental/spatial variability and prokaryotic community  $\beta$ NTI derived from Mantel testes with 999 permutations, r<sub>temporal</sub> correlation coefficients between environmental/spatial variability and prokaryotic community  $\beta$ NTI controlled by temporal variation derived from partial Mantel testes with 999 permutations. Data in bold indicate significant correlations ( $P < 0.05$ ). Refer to Table 1 and Table 2 for abbreviations.

phylogenetic diversity (faith\_pd) (Dataset S2), and the correlation coefficient between AN and the weighted UniFrac distance-based  $\beta$ -diversity was the highest (Table 2). The correlation between un\_ionN and  $\alpha/\beta$ -diversity was consistent with that of AN, and even with higher coefficients (Tables 2 & 4, Dataset S2), the effect of un\_ionN on the microbial community was rarely reported, since un\_ionN is a part of AN, and in this study, un\_ionN exhibited a different influence on the distribution of ssOTUs among seasons (Fig. S2) with AN, which may indicate that the different chemical forms on phylogenetic distribution are different, and more targeted research is needed to explain this phenomenon. There was a negative correlation between SRP and the prokaryotic  $\alpha$ -diversity in all samples, while the analysis based on a single sampling time showed that it was in August that the  $\alpha$ -diversity indexes (faith\_pd, richness, Shannon, and Pielou\_e) were all negatively correlated with SRP (Dataset S2), which indicated that it may be the heterogeneous input of phosphate in different stations in August that caused the variation. Gilbert et al.<sup>31</sup> reported that the most abundant phyla varied among different seasons in the Western English Channel and were associated with SRP. Here, we also found the importance of SRP in shaping the phylogenetic structure of the prokaryotic community. Phosphate has been identified to limit primary productivity<sup>22</sup>, and in this study, phosphate played an important role in structuring the dominant taxa enriched in Oct (ssOTUs in Oct), which was mainly composed of members from Gammaproteobacteria (Fig. 4c), whose variation was also reported to be significantly explained by the SRP variable in the marine environment<sup>23</sup>. COD was also found to be an important factor significantly correlated with  $\beta$ -diversity and  $\beta$ NTI (Tables 2, 3) in this study. COD serves as one of the comprehensive indexes of the relative content of organic matter, which reflects the degree of pollution by reducing substances in the water. In the coastal region of the East China Sea, it was also found that COD had a significant effect on the microbial community composition<sup>4</sup>. The COD with the highest value was observed in Oct in Bohai Bay in this study, which may be caused by the apoptosis of phytoplankton since *Chl\_a* was the lowest in Oct, which may contribute to the higher COD. In addition, the seasonal COD flux and tidal influence may also affect the transport processes of COD in the Bohai Bay area in different seasons<sup>32</sup>. DO and DIN had relatively lower but significant correlations with  $\alpha/\beta$ -diversity and  $\beta$ NTI than other parameters, but they were both reported as very important factors in shaping marine bacterial communities<sup>18,33–35</sup>. In conclusion, the coexistence of multiple seasonally varied environmental factors shows a significant association with the prokaryotic seasonal pattern of surface water in the study area.

According to the quantitative analysis of assembly processes among different months, it could be clearly seen that the selection processes accounted for most (> 80%) of the turnover of the community between any two sampling times, and combined with the  $\beta$ NTI value mostly above 2, it can be speculated that the “heterogeneous selection” processes dominated the temporal community variation (Fig. 5). The temporal effect was so strong that the spatial effect was masked by it during different sampling times. Recently, increasing evidence has shown that the spatial distributions of microbes are nonrandom and that distance decay of similarity occurs for microbial compositional structures<sup>36,37</sup>. Similar findings have been made in Bohai Bay as a result of our previous research within the study area, similar to this work<sup>18</sup>. To eliminate the influence of seasonal factors on spatial effects, we also tested the environmental-spatial effect on PCC variation for single season/month samples. According to

the VPA results, no factors could explain the community variation in May (Table 3), and the assembly process quantification results showed that the “undominated” processes accounted for 58.33% of the community turnover in May among the 12 sites, while the “selection process” accounted for 25.72%, which indicated that the communities in May were mostly governed by stochastic processes, and there may be unknown environmental factors that were not detected that accounted for the “selection process” parts (Fig. 5C). In Aug, the environmental and spatial factors did not show purely significant effects on the PCC variation, while their joint effects could explain 49.5% of the PCC variation (based on UUF distance) among the 28 sites in this sampling time (Table 3), and the dispersal limitation accounted for 66.41% of the community turnover in Aug (Fig. 5C), which suggested that spatial factors play an important role in determining the PCC variation in Aug, since in multivariate analysis, when using variation partitioning<sup>38</sup> to estimate the relative importance of environmental and spatial factors, the spatial factors often considered to infer the dispersal process in microbial community assembly in marine waters, which could be considered as spatial effects. In October, the VPA results based on unweighted UniFrac distances showed that the environmental and spatial factors could explain 34.7% of the PCC variation. Meanwhile, the quantitative analysis of assembly processes based on unweighted  $\beta$ NTI showed that the selection and dispersal process accounted for 36.6% of the community turnover, which was consistent with the VPA results (Table 3, Fig. 5C). From the above, we could conclude that environmental and spatial factors driving the assembly processes on the PCC variation were changed with the temporal shifts.

The detection of dominant groups within microbial assemblages may be of particular importance in ecological studies concerning biogeochemical cycles<sup>2</sup>, as microbes belonging to different groups can express different degrees of activity in a given ecosystem<sup>39</sup>. The groups of microorganisms that varied with the season variation were different in different studies. For example, Gammaproteobacteria in the SPOT observation was the highest in August and the lowest in May, Flavobacteria was highest in February and lowest in Oct, while Cyanobacteria was highest in November and lowest in April<sup>6</sup>. A similar study on the Qinhuangdao coast of the Bohai Sea found that Alphaproteobacteria and Firmicutes were the most abundant in July, while Cyanobacteria and Bacteroidetes were the most abundant in Oct<sup>5</sup>. In contrast, in this study, the abundance of Gammaproteobacteria was highest in Oct and lowest in May, that of Alphaproteobacteria and Firmicutes was highest in May, while that of Bacteroidetes was highest in August. In contrast, Cyanobacteria did not show significant differences at each sampling time during the three sampling times in different seasons (Table S2). The differences between studies in different regions indicate that seasonal variation in microorganisms is not consistent across the region, even in different regions of the same ocean. In addition to differences in detection methods (qPCR/ARISA/Illumina MiSeq), the differences in the results further indicate that except for seasonal factors, other local environmental factors may also play an important role in determining prokaryotic community patterns. Further study is needed to focus on seasonal changes in microbial communities in different marine environments.

This study paid more attention to OTUs that varied dramatically among the detected months with the aim of finding their correlation with the local environmental factors, which may provide some clues for the prediction of how the microbial community reacts to changes in the marine bay environment. We found that the ssOTUs naturally clustered as different modules in the network analysis (Fig. 4A), and the cumulative relative abundance of the ssOTUs in each module all showed definite seasonal bias (Fig. 4B). Moreover, the ssOTUs had a significant correlation with the environmental factors (Fig. S2, Temperature, AN, un\_ionN, SRP, pH), which was consistent with the driving factors related to the PCC temporal variation obtained from the Mantel test based on  $\beta$ -diversity. These results indicated that the seasonally varied environmental factors definitely affected the variation in prokaryotic communities at the OTU level. Many studies have aimed to identify the specific taxonomic groups that drive seasonal variation, and it was found that nonseasonal taxa contain seasonal OTUs<sup>6</sup>, which calls for finer taxonomic groups to be analysed since they may adapt to different ecological niches using different metabolic genes. In this study, the analysis of ssOTUs could reflect the response of the prokaryotic communities to temporal environmental changes more elaborately and comprehensively. It was found that some ssOTUs within the phyla to which they belonged (Verrucomicrobia, Betaproteobacteria, Crenarchaeota) did not show significant temporal variation, and most of the ssOTUs in different seasons came from the same phylum. This suggested that it may not be enough to analyse the community at only the phylum level but may be necessary to pay more attention to finer levels.

Temporal functional gene variation has rarely been studied in seasonal research on marine environments, such as in Bohai Bay. In this study, PICRUSt2 was used to predict the approximate functional potential of the prokaryotic community. The accuracy of PICRUSt generally increased with decreasing NSTI score, and reliable results were generated from a dataset of soil samples with a mean NSTI score of 0.17. In this study, the mean NSTI score was 0.058, suggesting that the results are credible and could accurately reflect the variation trend of the targeted function. Based on the PICRUSt2 results, the nitrogen fixation and sulfur oxidation functions were mainly enriched in August, while the denitrification and sulfite or sulfate reduction functions were mainly enriched in May and October (Fig. 6). The variation in nitrogen metabolism during different seasons in this study was consistent with a study of oceanic mangrove habitats, where the rate of nitrogen fixation in the wet season was higher than that in the dry season, while the potential denitrification rates were higher during the dry season than during the wet season<sup>40</sup>. The variation in sulfur metabolism in different seasons in surface seawater has rarely been reported, since it is often reported to occur in anaerobic environments, and most studies are concerned with the variation in sulfur metabolism in lakes and marine sediments<sup>16,41</sup>. In this study, the nitrogen fixation and sulfur oxidation functions were predicted to be enriched in August, which may be because rainfall in the wet season (Fig. S6) is often accompanied by periodic high tides, which bring abundant dissolved nutrients, such as SRP, nitrate, and ammonia nitrogen, which could effectively improve the availability of nitrogen or sulfur, promoting the enrichment of nitrogen fixation and sulfur oxidation. However, the anaerobic functions mediated by sulfate-reducing bacteria and denitrifying bacteria that were enriched in May and Oct should be carefully interpreted. This phenomenon may be explained by the relative abundance of the prokaryotic taxa

that contributed to the denitrification and sulfate reduction functions. Prokaryotic taxa contributing to these two functions at the species level showed a very low relative abundance of minus than 0.01 for most of them (Fig. S5), except for taxa belonging to Rhodobacteraceae, which contained members with aerobic denitrification abilities, such as *Paracoccus*, a representative aerobic denitrifying bacterium that was commonly found to be dominant in sulfate reduction bacteria (SBR) during the denitrification process<sup>42</sup>. The contributed taxa varied among the three sampling times, resulting in the corresponding functional variation, while this was not to say that the detected function played a key role in the corresponding period; it could just reflect the relative change in a certain function during the sampling period in different seasons. However, the distinct variation in nitrogen and sulfur metabolism-related genes distributed in different sampling months reminds us to pay more attention to the nitrogen and sulfur cycles in surface seawaters in future research.

The linkage between nitrogen and sulfur metabolism is also worth attention. In this study, both the denitrification and sulfate reduction potential functions were predicted to be enriched in May and Oct, which may indicate that these two processes are interwoven with each other in the surface water of Bohai Bay. These predicted denitrifying members of Alphaproteobacteria and Gammaproteobacteria were most negatively correlated with nitrogen nutrients (Dataset S9), which may suggest the strengthened consumption of nitrogen salts by these taxa in October and May. Simultaneously, the sulfite and sulfate reduction members of Deltaproteobacteria were most positively correlated with  $\text{un\_ionN}$ , which may indicate the relevance between sulfate reduction and nitrogen metabolism. These findings may guide further study of nitrogen and sulfur metabolism in Bohai Bay, and more reliable analytical techniques, such as metagenomic and meta-transcriptomic sequencing, will be needed in future work.

Finally, due to limited sampling, samples from the cold season were not collected, while the aim of this study was to explore the spatial and temporal effects on the distribution of prokaryotic communities. The results clearly showed strong “heterogeneous selection” processes affecting the PCC temporal variation during the study period. The VPA and quantitative analysis of the assembly processes showed that environmental and spatial factors driving the assembly processes on the PCC variation were changed with the temporal shifts. There were distinct seasonal features at the phylum and OTU levels (ssOTUs), while the ssOTUs exhibited different seasonal attributions under the same phylum, which calls for finer-level analysis to reflect the temporal variation more strictly. Moreover, functional prediction showed that nitrogen and sulfur metabolism were significantly shifted during the temporal variation, which needs further study for nitrogen and sulfur metabolism in Bohai Bay.

## Materials and methods

**Study area, sampling, and physicochemical analyses of seawater samples.** We collected 188 (biological replicates included) surface seawater samples from 12 to 28 stations (Fig. S1), which covered the north and south coasts and the offshore area of Bohai Bay. Among the sampling stations, TJ19 was located nearest the Bei Pai River mouth, TJ12 was located near the Du Liu Jian River in the southern part, TJ04 was located near the Yong Ding Xin River in the northern part, TJ01 was located near Da Sheng Tang village, TJ02 was located near Ma Zu park, TJ21 was located near Dongjiang Bay, which are all popular tourist attractions in the Tianjin coastal region, TJ13, TJ09, TJ20, TJ24 and TJ30 were located at a middle distance off the shore, and the most remote station, TJ14, was located as far as 110 km off the shore, which was more influenced by ocean currents from the Bohai Sea than other stations. Surface water samples at 0.5 m depth were collected at each site during 16–18 May, 16–18 August (Aug) and 17–19 October (Oct) in 2016. In May, samples from 12 stations were collected, and in August and October, samples from 28 stations were collected. In Aug and Oct, each site was sampled with three replicates and in May with two replicates. The detailed sampling information is listed in Dataset S1. All sampling was undertaken during high tide conditions. Approximately 500 ml of the water sample of each sample was filtered onto a 0.2  $\mu\text{m}$  polycarbonate membrane (Millipore, USA) and stored at  $-20^\circ\text{C}$  prior to DNA extraction. Environmental DNA on the polycarbonate membrane was extracted using a PowerWater DNA Isolation Kit (MO BIO, USA) according to the manufacturer's instructions. Sampling environmental parameters, including water temperature, depth and transparency, were recorded through field measurements. The pH values, DO, salinity and conductivity were measured using a probe (Multi 3430, WTW, Germany). The concentrations of soluble reactive phosphate (SRP), ammonium nitrogen (AN), nitrite, nitrate, dissolved inorganic nitrogen (DIN) and COD were analysed following the standard methods of the specification for marine monitoring of seawater analysis (GB 17378.4-2007). Chlorophyll *a* (*Chl\_a*) was determined in water samples in accordance with the “Specification for oceanographic survey” (GB/T 12763.4-2007). Nonionic ammonia ( $\text{un\_ionN}$ ) was determined in accordance with the “Surface water quality standards” (GB 3838-88). The geographic coordinates and physicochemical parameters of the water samples are listed in Dataset S1. Student's *t* test was used to compare the environmental difference between any two time points. Principal component analysis (PCA) was performed to reveal the relationships between the environmental parameters and the distribution of the sampling stations and time in an ordination plot.

**DNA extraction, 16S rRNA gene amplification, and Illumina sequencing.** The DNA extracts were quantified using a spectrophotometer (NanoDrop ND-2000, USA) and subsequently submitted to Novogene Co. Beijing for 16S rRNA gene amplification and sequencing. Briefly, the 16S rRNA gene V4 region was amplified using dual-indexed bacterial/archaeal primers 515f. (GTG YCA GCM GCC GCG GTA A) and 806r (GGA CTA CNV GGG TWT CTA AT) containing adaptor sequences for the Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA) using the 2 × 250PE rapid run mode. Details of the sequencing libraries can be found in previous work<sup>18</sup>.



**16S rRNA gene amplicon processing and analysis.** Sequences were demultiplexed in Quantitative Insights Into Microbial Ecology version 2 (QIIME2). The raw sequences with forwards and reverse reads were imported through *time QIIME tools import* with PairedEndFastqManifestPhred33V2 as input format, and amplicon sequence variants were generated using DADA2 version 2019.7.0 with *time QIIME dada2 denoise-paired* script to remove sequences with quality scores < 20 and chimeric sequences<sup>43</sup>. FeatureData[Sequence] were obtained after quality control with DADA2, which is the equivalent of the QIIME 1 representative sequences file, and which are the equivalent of 100% OTUs from QIIME 1 with more accurate estimates of diversity and taxonomic composition, which are generally referred to as amplicon sequence variants (ASVs). The phylogenetic tree was generated based on the representative sequences using align-to-tree-mafft-fasttree from the Q2-Phylogeny plugin. Taxonomy was assigned using the q2-feature-classifier command with default parameters in QIIME2, and sequences were matched against the Greengenes 13\_8 database<sup>44</sup>.

Prokaryotic  $\alpha$ -diversity and  $\beta$ -diversity estimates were calculated by even rarefaction at 26,000 reads per sample, with multiple indexes (Faith Phylogenetic Diversity metric (Faith\_pd), the quantification of taxon richness through phylogenetic tree units (OTU\_richness), Shannon and evenness (pielou\_e) and weighted (WUF)/unweighted UniFrac (UUF) distances between samples. The differences in  $\alpha$ -diversity indexes and taxa abundances at the phylum level among different sampling times were analysed by one-way ANOVA with alpha levels obtained by a two-tailed test. Pearson's correlation coefficients between geo-physical-chemical variables and  $\alpha$ -diversity indexes were calculated to measure the relationships between them. We used a principal coordinates analysis (PCoA) plot based on WUF and UUF distances to visualize the sample clusters and dissimilarity in prokaryotic community composition between sampling times and subsequently tested the significance of pairwise dissimilarity in community composition between sampling times with analysis of similarity (ANOSIM) and permutational multivariate analysis of variance (PERMANOVA). Simple and partial Mantel tests were used to test the correlations of geographic distance and environmental parameters with prokaryotic  $\beta$ -diversity (variation in community composition). Variance partitioning analysis (VPA) was conducted on all samples or single-time sampling using the WUF/UUF distances as response variables and using environmental factors, linear trends, and principle coordinates of neighbor matrices (PCNM) as the explanatory variables to investigate the relative contributions of spatial and environmental parameters to the variation in PCC, with the same analysis methods as those used in a former study in Bohai Bay<sup>18</sup>. The spatial parameters included linear trends and PCNM, and the environmental factors included all the physical-chemical parameters detected in Bohai Bay seawater samples.

The OTUs responsible for the differences among the sampling times were identified by correlation-based indicator species analysis using the function *multipatt()* in the package "indicspecies"<sup>45</sup> and likelihood ratio tests using the function *DGEList()* in the package "edgeR"<sup>46</sup>. Finally, the OTUs confirmed by both approaches were defined as seasonal sensitive OTUs (ssOTUs). The cooccurrence network was performed on the OTUs with a relative abundance of > 0.01%. Then, the abundances of screened OTUs were normalized by the "trimmed means of M" (TMM) method and expressed as relative abundance per counts million (CPM) using the function *cpm()* in the package "edgeR". The TMM-normalized CPM counts and calculated Spearman rank correlations ( $\rho > 0.7$ ,  $p < 0.001$ ) between OTUs were utilized to construct a meta co-occurrence network. To further investigate the relationship of ssOTUs and the community structure of different sampling times, we identified net-work modules, which are substructures of nodes with a higher density of edges within groups than between them. For this, we utilized the greedy optimization of modularity algorithm<sup>47</sup> as implemented in the package "igraph". Modules with the ssOTUs were further analysed to calculate the cumulative relative abundance using the TMM method; simultaneously, the regression analysis between the module abundance and environmental factors was also analysed through  $R$ <sup>48</sup>.

The functional potential of the prokaryotic community was predicted by PICRUSt2 (v2.1.4 beta)<sup>49</sup> according to the guidelines with the unrarefied OTU abundance table as the input. The relative predicted abundance of KEGG pathways was calculated by dividing the abundance of each pathway by the sum of all pathway abundances per sample. The relative contribution of each OTU to the predicted pathways was calculated by dividing the contribution of each OTU by the sum of all contributions per sample. The significantly discriminant energy metabolism gene ontologies related to nitrogen and sulfur metabolism in each sampling time were determined using LDA effect size<sup>50</sup>, which employs the factorial Kruskal–Wallis sum-rank test ( $\alpha = 0.05$ ) to identify genes with significantly different relative abundances between sampling times (using one-against-all comparisons). The taxonomy assignment of the predicted genes associated with nitrogen or sulfur metabolism with relative abundance significantly different in each season was plotted using Heml 1.0<sup>51</sup>.

**Analysis of assembly processes of prokaryotic communities.** The  $\beta$ -nearest taxon index ( $\beta$ NTI)<sup>12</sup> and Bray–Curtis-based Raup–Crick (RCbray)<sup>52</sup> were used to quantify community assembly processes and identify features that impose them. Based on the random shuffling of OTU labels across the tips of the phylogeny, the  $\beta$ NTI was computed as the number of standard deviations that observed  $\beta$ -mean nearest taxon distance ( $\beta$ MNTD) values deviated from the mean of the null distribution (999 null replicates)<sup>12</sup>. If the observed  $\beta$ MNTD is significantly greater or less than the null expected phylogenetic turnover ( $\beta$ NTI > +2 or < -2), we can infer that "heterogeneous selection" or "homogeneous selection" (under heterogeneous or homogeneous environmental conditions) governs phylogenetic difference or similarity between communities, respectively<sup>53,54</sup>. If  $|\beta$ NTI| < 2, stochastic processes may play an important role in structuring microbial communities<sup>12,17</sup>. The second step of the quantification was to use the modified Raup–Crick metric<sup>55</sup> based on Bray–Curtis dissimilarity to evaluate the standardized deviation of observed OTU compositional turnover from the null expectation<sup>52</sup>. When  $|\beta$ NTI| < 2, if the observed Bray–Curtis dissimilarity is significantly greater or less than the null expected compositional turnover (RCbray > +0.95 or < -0.95), we can infer that "dispersal limitation" or "homogenizing dispersal" governs compositional difference or similarity between communities, respectively. RCbray values



between  $-0.95$  and  $+0.95$  infer that the compositional turnover between a given pair of communities is “undominated” by any single process<sup>54</sup>. Subsequently, we calculated the percentages of each process in all pairwise comparisons according to the  $\beta$ NTI and RCbray values.

The boxplot or bar plot in Figs. 2, 3, 5 and 6 were made through the ImageGP web server<sup>39</sup>.

## Data availability

The raw data of sequencing samples have been deposited under the sequence read archive (SRA) section of NCBI under the following accession number: PRJNA645045, which are available from the SRA Run Selector: [https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA645045&o=acc\\_s%3Aa](https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA645045&o=acc_s%3Aa).

Received: 16 May 2022; Accepted: 22 November 2022

Published online: 28 November 2022

## References

- Celussi, M., Bussani, A., Cataletto, B. & Del, N. P. Assemblages' structure and activity of bacterioplankton in northern Adriatic Sea surface waters: A 3-year case study. *FEMS Microbiol. Ecol.* **75**, 77. <https://doi.org/10.1111/j.1574-6941.2010.00997.x> (2011).
- Castle, D. & Kirchman, D. L. Composition of estuarine bacterial communities assessed by denaturing gradient gel electrophoresis and fluorescence in situ hybridization. *Limnol. Oceanogr. Methods* **2**, 303–314. <https://doi.org/10.4319/lom.2004.2.303> (2004).
- Fuhrman, J. A. *et al.* Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13104–13109. <https://doi.org/10.1073/pnas.0602399103> (2006).
- Wang, K. *et al.* Bacterial biogeography in the coastal waters of northern Zhejiang, East China Sea is highly controlled by spatially structured environmental gradients. *Environ. Microbiol.* **17**, 3898–3913. <https://doi.org/10.1111/1462-2920.12884> (2015).
- He, Y. D. *et al.* Distinct seasonal patterns of bacterioplankton abundance and dominance of phyla alpha-proteobacteria and cyanobacteria in Qinhuangdao coastal waters off the Bohai Sea. *Front. Microbiol.* **8**, 1579. <https://doi.org/10.3389/fmicb.2017.01579> (2017).
- Cram, J. A. *et al.* Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J.* **9**, 563–580. <https://doi.org/10.1038/ismej.2014.153> (2015).
- Xiong, J. B. *et al.* Biogeography of the sediment bacterial community responds to a nitrogen pollution gradient in the East China Sea. *Appl. Environ. Microbiol.* **80**, 1919–1925. <https://doi.org/10.1128/aem.03731-13> (2014).
- Du, J. K. *et al.* Temporal and spatial diversity of bacterial communities in coastal waters of the South China Sea. *PLoS ONE* **8**, e66968. <https://doi.org/10.1371/journal.pone.0066968> (2013).
- Suh, S. S. *et al.* Seasonal dynamics of marine microbial community in the South Sea of Korea. *PLoS ONE* **10**, e0131633. <https://doi.org/10.1371/journal.pone.0131633> (2015).
- Zhang, R., Lau, S. C., Ki, J. S., Thiagarajan, V. & Qian, P. Y. Response of bacterioplankton community structures to hydrological conditions and anthropogenic pollution in contrasting subtropical environments. *FEMS Microbiol. Ecol.* **69**, 449–460. <https://doi.org/10.1111/j.1574-6941.2009.00726.x> (2009).
- Reza, M. S. *et al.* Basin-scale seasonal changes in marine free-living bacterioplankton community in the Ofunato Bay. *Gene* **665**, 185–191. <https://doi.org/10.1016/j.gene.2018.04.074> (2018).
- Stegen, J. C., Lin, X., Konopka, A. E. & Fredrickson, J. K. Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* **6**, 1653–1664. <https://doi.org/10.1038/ismej.2012.22> (2012).
- Dini-Andreote, F., Stegen, J. C., van Elsland, J. D. & Salles, J. F. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E1326–1332. <https://doi.org/10.1073/pnas.1414261112> (2015).
- Vellend, M. Conceptual synthesis in community ecology. *Q. Rev. Biol.* **85**, 183–206. <https://doi.org/10.1086/652373> (2010).
- Chase, J. M. & Myers, J. A. Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **366**, 2351–2363. <https://doi.org/10.1098/rstb.2011.0063> (2011).
- Wang, K. *et al.* Regional variations in the diversity and predicted metabolic potential of benthic prokaryotes in coastal northern Zhejiang, East China Sea. *Sci. Rep.* **6**, 38709. <https://doi.org/10.1038/srep38709> (2016).
- Zhang, H. J. *et al.* Microbial community dynamics and assembly follow trajectories of an early-spring diatom bloom in a semi-enclosed bay. *Appl. Environ. Microbiol.* **84**, e01000–01018. <https://doi.org/10.1128/aem.01000-18> (2018).
- Zhao, W. *et al.* Bacterioplankton community variation in Bohai Bay (China) is explained by joint effects of environmental and spatial factors. *Microbiologyopen* **9**, e997. <https://doi.org/10.1002/mbo3.997> (2020).
- Guo, L., Sun, M., Wang, X. & Lan, H. Statistical analysis of the characteristics difference of rainfall over Bohai Bay and the land. *Clim. Change Res. Lett.* **8**, 835–844. <https://doi.org/10.12677/CCRL.2019.86092> (2019).
- Witt, V., Wild, C., Anthony, K. R., Diaz-Pulido, G. & Uthicke, S. Effects of ocean acidification on microbial community composition of, and oxygen fluxes through, biofilms from the Great Barrier Reef. *Environ. Microbiol.* **13**, 2976–2989. <https://doi.org/10.1111/j.1462-2920.2011.02571.x> (2011).
- Shore, A., Day, R. D., Stewart, J. A. & Burge, C. A. Dichotomy between regulation of coral bacterial communities and calcification physiology under ocean acidification conditions. *Appl. Environ. Microbiol.* <https://doi.org/10.1128/aem.02189-20> (2021).
- Krom, M., Kress, N. & Gordon, S. Phosphorus limitation of primary productivity in the eastern Mediterranean Sea. *Limnol. Oceanogr.* **36**, 424–432. <https://doi.org/10.4319/lo.1991.36.3.0424> (1991).
- Gilbert, J. A. *et al.* The seasonal structure of microbial communities in the Western English Channel. *Environ. Microbiol.* **11**, 3132–3139. <https://doi.org/10.1111/j.1462-2920.2009.02017.x> (2009).
- Arp, D. J., Stein, L. Y. Metabolism of inorganic N compounds by ammonia-oxidizing bacteria. *Crit Rev Biochem Mol Biol.* **38**(6), 471–495. <https://doi.org/10.1080/10409230390267446> (2003).
- Apple, J. K., Giorgio, P. A. D. & Kemp, W. M. Temperature regulation of bacterial production, respiration, and growth efficiency in a temperate salt-marsh estuary. *Aquat. Microb. Ecol.* **43**, 243–254. <https://doi.org/10.3354/ame043243> (2006).
- Fonseca, F., Cerqueira, R. & Fuentes, J. Impact of ocean acidification on the intestinal microbiota of the marine Sea Bream (*Sparus aurata* L.). *Front. Physiol.* **10**, 1446. <https://doi.org/10.3389/fphys.2019.01446> (2019).
- Lindh, M. V. *et al.* Consequences of increased temperature and acidification on bacterioplankton community composition during a mesocosm spring bloom in the Baltic Sea. *Environ. Microbiol. Rep.* **5**, 252–262. <https://doi.org/10.1111/1758-2229.12009> (2013).
- Meron, D. *et al.* The impact of reduced pH on the microbial community of the coral *Acropora eurystroma*. *ISME J.* **5**, 51–60. <https://doi.org/10.1038/ismej.2010.102> (2011).
- Hassenruck, C., Hofmann, L. C., Bischof, K. & Ramette, A. Seagrass biofilm communities at a naturally CO<sub>2</sub>-rich vent. *Environ. Microbiol. Rep.* **7**, 516–525. <https://doi.org/10.1111/1758-2229.12282> (2015).
- Richa, K. *et al.* Distribution, community composition, and potential metabolic activity of bacterioplankton in an urbanized Mediterranean Sea coastal zone. *Appl. Environ. Microbiol.* **83**, e00494–e1417. <https://doi.org/10.1128/aem.00494-17> (2017).

31. Gilbert, J. A., Field, D., Swift, P., Newbold, L. & Joint, I. The seasonal structure of microbial communities in the Western English Channel. *Environ. Microbiol.* **11**, 3132–3139. <https://doi.org/10.1111/j.1462-2920.2009.02017.x> (2010).
32. Jie, G. U., Qian, C. R., Liang, H. D., Kuang, C. P. & Zuo, L. M. Seasonal and spatial distribution of chemical oxygen demand(COD) in Caofeidian sea area. *Mar. Environ. Sci.* <https://doi.org/10.13634/j.cnki.mes.2017.04.015> (2017).
33. Long, A. M., Jurgensen, S. K., Petchel, A. R., Savoie, E. R. & Brum, J. R. Microbial ecology of oxygen minimum zones amidst ocean deoxygenation. *Front. Microbiol.* **12**, 748961. <https://doi.org/10.3389/fmicb.2021.748961> (2021).
34. Merlo, C. *et al.* Changes in the bacterial community composition of different habitats along a polluted river (Suquia River, Cordoba, Argentina). *Ecol. Aust.* **27**, 72–84 (2017).
35. Caffrey, J. M., Bano, N., Kalanetra, K. & Hollibaugh, J. T. Ammonia oxidation and ammonia-oxidizing bacteria and archaea from estuaries with differing histories of hypoxia. *ISME J.* **1**, 660–662. <https://doi.org/10.1038/ismej.2007.79> (2007).
36. Martiny, J. B. *et al.* Microbial biogeography: Putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**, 102–112. <https://doi.org/10.1038/nrmicro1341> (2006).
37. Wang, J. *et al.* Phylogenetic beta diversity in bacterial assemblages across ecosystems: Deterministic versus stochastic processes. *ISME J.* **7**, 1310–1321. <https://doi.org/10.1038/ismej.2013.30> (2013).
38. Legendre, P. *et al.* Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecol. Monogr.* **75**, 435–450 (2005).
39. Matthew, T. C. & David, L. K. Part 1 || contribution of major bacterial groups to bacterial biomass production (thymidine and leucine incorporation) in the delaware estuary. *Limnol. Oceanogr.* **48**, 168–178 (2003).
40. Lee, R. Y. & Joye, S. B. Seasonal patterns of nitrogen fixation and denitrification in oceanic mangrove habitats. *Mar. Ecol. Prog. Ser.* **307**, 127–141. <https://doi.org/10.3354/meps307127> (2006).
41. Diao, M., Huisman, J. & Muyzer, G. Spatio-temporal dynamics of sulfur bacteria during oxic-anoxic regime shifts in a seasonally stratified lake. *FEMS Microbiol. Ecol.* **94**, fty040. <https://doi.org/10.1093/femsec/fty040> (2018).
42. Kwon, J. H., Park, H. J. & Lee, Y. Y. Evaluation of denitrification performance and bacterial community of a sequencing batch reactor under intermittent aeration. *J. Environ. Sci. Health Part A Toxic Hazard. Subst. Environ. Eng.* **55**, 179–192. <https://doi.org/10.1080/10934529.2019.1681220> (2020).
43. Bolyen, E. R. J. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857. <https://doi.org/10.1038/s41587-019-0209-9> (2019).
44. McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618. <https://doi.org/10.1038/ismej.2011.139> (2012).
45. De Cáceres, M., Legendre, P. & Moretti, M. Improving indicator species analysis by combining groups of sites. *Oikos* **119**, 1674–1684. <https://doi.org/10.1111/j.1600-0706.2010.18334.x> (2010).
46. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxf. Engl.)* **26**, 139–140. <https://doi.org/10.1093/bioinformatics/btp616> (2010).
47. Clauset, A., Newman, M. E. & Moore, C. Finding community structure in very large networks. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **70**, 066111. <https://doi.org/10.1103/PhysRevE.70.066111> (2004).
48. Hartman, K. *et al.* Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. *Microbiome* **6**, 14–14. <https://doi.org/10.1186/s40168-017-0389-9> (2018).
49. Douglas, G. M. *et al.* PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **38**, 685–688. <https://doi.org/10.1038/s41587-020-0548-6> (2020).
50. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60. <https://doi.org/10.1186/gb-2011-12-6-r60> (2011).
51. Deng, W. K., Wang, Y. B., Liu, Z. X., Cheng, H. & Xue, Y. HemI: A toolkit for illustrating heatmaps. *PLoS ONE* **9**, e111988. <https://doi.org/10.1371/journal.pone.0111988> (2014).
52. Stegen, J. C. *et al.* Quantifying community assembly processes and identifying features that impose them. *ISME J.* **7**, 2069–2079 (2013).
53. Zhou, J. & Ning, D. Stochastic community assembly: Does it matter in microbial ecology?. *Microbiol. Mol. Biol. Rev. MMBR* <https://doi.org/10.1128/mmb.00002-17> (2017).
54. Stegen, J. C., Lin, X., Fredrickson, J. K. & Konopka, A. E. Estimating and mapping ecological processes influencing microbial community assembly. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2015.00370> (2015).
55. Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M. & Inouye, B. D. Using null models to disentangle variation in community dissimilarity from variation in  $\alpha$ -diversity. *Ecosphere* **2**, art24. <https://doi.org/10.1890/ES10-00117.1> (2011).

## Acknowledgements

We would like to thank the staff of the Tianjin Survey and Design Institute for Water Transport Engineering and Tianjin Key Laboratory of Marine Resources and Chemistry, Tianjin University of Science and Technology, for supporting our sampling. This work was supported by the National Key R&D Program of China (grant 2018YFA0901200), the Science and Technology Support Program of Tianjin Municipality (grant 18ZXSZSF00100, 19YFZCSF00750), the General Program of National Natural Science Foundation of China (grant 41977200), the Natural Science Foundation of Tianjin (grant 20JCYBJC01220), and the Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (Grant TSBICIP-IJCP-001).

## Author contributions

W.Z. conceived the manuscript, performed the data analysis and wrote the manuscript, L.S. conducted the experiments; S.X. and X.Z. contributed the sampling; L.M., W.Z. and Y.H. contributed the seawater quality analysis; X.W., J.W. and Z.H. contributed the manuscript revision.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-24892-1>.

**Correspondence** and requests for materials should be addressed to Z.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022