Composition and predictive functional analysis of bacterial communities in the surface seawater of the Changjiang Estuary

Dong-Mei Wu, Jian-Xin Wang, Xiao-Hui Liu, Ying-Ping Fan, Ran Jiang, Ming-Hua Liu, Shuai-Bing Wang, Xue-Zhu Liu

1 Marine Science and Technology College, Zhejiang Ocean University, Zhoushan, China

Corresponding Author: Xue-Zhu Liu
Email address: liuxuezhu@zjou.edu.cn

The objective of this study was to characterize the structure and function of microbial communities in surface seawater from the Changjiang Estuary and adjacent areas, China. Sample water was collected at 12 sites and environmental parameters were measured. Community structure was analyzed using high-throughput sequencing of 16S rDNA genes. Predictive metagenomic approach was used to predict the function of bacterial communities. Result showed that sample site A0102 had the highest bacterial abundance and diversity. The heatmap indicated that different samples could be clustered into six groups. Phylogenetic analysis showed that Proteobacteria was the predominant phylum in all samples, followed by Bacteroidetes and Actinobacteria. Alphaproteobacteria and Gammaproteobacteria were the dominant classes. The analysis of predictive metagenomic showed carbon fixation pathways in prokaryotes, nitrogen metabolism, carbon fixation in photosynthetic organisms, photosynthesis and polycyclic aromatic hydrocarbon degradation were enriched in all samples. Redundancy analysis (RDA) identified that dissolved oxygen (DO) and PO$_4^{3-}$ concentration had positive correlations with the bacterial communities while chemical oxygen demand (COD), dissolved oxygen (DO) and PO$_4^{3-}$ concentration were significantly associated with microbial functional diversity. This study adds to our knowledge of functional and taxonomic composition of microbial communities.
Composition and predictive functional analysis of bacterial communities in the surface seawater of the Changjiang Estuary

Dong-Mei Wu*, Jian-Xin Wang#, Xiao-Hui Liu, Ying-Ping Fan, Ran Jiang, Ming-Hua Liu, Shuai-Bing Wang, Xue-Zhu Liu*

Marine Microorganism Ecological & Application Lab, Zhejiang Ocean University, 1 South Haida Road, Zhejiang 316022, China.

*These authors contributed equally to this study and share the first authorship.

*Correspondence author. Tel.: +86 13575617505; fax: +86 580-8180982; E-mail: liuxuezhu@zjou.edu.cn

KEY WORDS: Bacterial community structure; 16S rDNA; Environmental factors; PICRUSt

ABSTRACT

The objective of this study was to characterize the structure and function of microbial communities in surface seawater from the Changjiang Estuary and adjacent areas, China. Sample water was collected at 12 sites and environmental parameters were measured. Community structure was analyzed using high-throughput sequencing of 16S rDNA genes. Predictive metagenomic approach was used to predict the function of bacterial communities. Result showed that sample site A0102 had the highest bacterial abundance and diversity. The heatmap indicated that different samples could be clustered into six groups. Phylogenetic analysis showed that Proteobacteria was the predominant phylum in all samples, followed by Bacteroidetes and Actinobacteria. Alphaproteobacteria and Gammaproteobacteria were the dominant classes. The
analysis of predictive metagenomic showed carbon fixation pathways in prokaryotes, nitrogen metabolism, carbon fixation in photosynthetic organisms, photosynthesis and polycyclic aromatic hydrocarbon degradation were enriched in all samples. Redundancy analysis (RDA) identified that dissolved oxygen (DO) and $\text{PO}_4^{3-}$ concentration had positive correlations with the bacterial communities while chemical oxygen demand (COD), dissolved oxygen (DO) and $\text{PO}_4^{3-}$ concentration were significantly associated with microbial functional diversity. This study adds to our knowledge of functional and taxonomic composition of microbial communities.

**INTRODUCTION**

Bacterioplankton communities are an important microorganisms in the marine ecosystems, which greatly affect material cycling, energy flow and the ocean food web. Microorganisms in aquatic ecosystems are very sensitive to changes in environmental conditions and thus bacterial community composition can act as an environmental indicator (Paerl et al. 2003). Determining physicochemical factors and bacterial community structures can increase our understanding of microbial ecology. To unravel the functional potential of bacteria is beneficial for understanding their roles in biogeochemical cycling.

The Changjiang Estuary, also called the Yangtze River. It located offshore from the mouth of the Changjiang River (Chen et al. 1999). Because of the mixture of Changjiang Dulited Water with the Taiwan Warm Current (TWC), this region is extremely complicated and dynamic (Jiao et al. 2007; Zhang et al. 1999). Many studies on bacterial diversity in the Changjiang Estuary have focused on ammonia-oxidizing bacteria (AOB), For example, molecular biological techniques were used to analyze the community structure and diversity of AOB in Changjiang Estuary sediments and adjacent waters in the East China Sea (Chen et al. 2014). Liu reported a relationship between bacterial abundance and concentrations of phosphate
in the Changjiang River (Liu et al. 2009). Sala, Vieira and colleagues studied the spatial
distribution of bacterial communities (Sala et al. 2008; Vieira et al. 2008). However, there is
relatively little information on the diversity and abundance of the whole bacterial population in
surface seawater of the Changjiang Estuary and adjacent areas. A growing number of studies
have focused on the functional potential of bacteria in marine sediments (Graves et al. 2016;
Kirchman et al. 2015). Information on the function of microbial communities in the surface
seawater are poorly understood.

In this study, we used high-throughput sequencing technology targeted to 16S rDNA
genes to analyze bacterial diversity and used predicted metagenomic analysis to compare
microbial functions. Meanwhile, we estimated the bacterial biomass based on 4,6-diamidino-2-
phenylindole (DAPI) fluorescence direct counts, and used clustering analyses to assess the
correlations between bacterial community structure, functions and environmental factors. Our
research is of benefit for understanding the bacterial abundance, diversity, functions and
distribution in the Changjiang Estuary and adjacent waters.

MATERIALS AND METHODS

Sampling areas and sampling

In July 2015, samples were collected using an SBE 32 sampler (Sea-Bird Electronics,
Washington, USA) at 2 m depth in surface seawater from 12 sites in the Changjiang Estuary and
adjacent areas (Figure 1). Table 1 lists the latitude and longitude of the sampling sites. Samples
for DAPI fluorescence examination were collected in 10-mL sterile cryopreservation tubes,
preserved with buffered glutaraldehyde (final concentration 1%), and stored in the dark at
ambient temperature for 15 min. Subsequently the samples were stored in airtight plastic bottles
at −20°C for the duration of the cruise, and at −80°C after returning to the laboratory. Samples
for DNA isolation were collected using a vacuum pump suction filter with a vacuum of 20 kPa (Porter & Feig 1980). One liter samples were filtered first through 3-μm then 0.22-μm pore size polycarbonate nucleopore membranes (Merck Millipore Ltd., USA). The 0.22-μm filters were preserved in 5-ml sterile cryopreservation tubes at −20°C during the cruise and at −80°C after returning to the laboratory.

**Environmental parameters**

Samples were pretreated according to specifications for marine monitoring (National Standards of People’s Republic of China, GB 17378.5, 2007) before chemical parameter analysis (Heijs et al. 2008). PO$_4^{3-}$ was measured using a QuAAtro continuous flow analyzer (SEAL Analytical, Hamburg, Germany). Dissolved oxygen (DO), NO$_2$ and NH$_4^+$ levels were measured using a spectrophotometer (752 UV/visible spectrophotometer, Shanghai-Hengping, China). Chemical oxygen demand (COD) was measured using the alkaline potassium permanganate.

**DAPI fluorescence direct counts**

Seawater samples were stained with DAPI by adding 1 mL of pretreated sample to 1 mL of 20 μg/mL DAPI. Staining was performed in a darkened room over 30 min with occasional swirling of the reaction tube. The mixed solution was then passed through pre-wetted black polycarbonate nucleopore membrane filters (pore size 0.2 μm) using a 5-mL syringe fitted with a needle. The membrane was observed under a fluorescence microscope using a flat-field 100× oil immersion lens, and a minimum of 30 cells per filter were counted in a minimum of 20 fields of view. Bacterial density in the original sample was calculated by using the formula (cell/mL) = (N × At)/(Ag × Vf), where N is the number of cells counted, At is the effective area of the filter (in mm$^2$ or μm$^2$), Ag is the area of the counting grid (in mm$^2$ or μm$^2$), and Vf is the volume of diluted sample filtered (in mL)(Jr & Pratt 1994).
DNA extraction and PCR amplification

Bacterial 16S rRNA genes were analyzed to determine the bacterial community structure and diversity. Genomic DNA was extracted using the FastPrep®-24 rapid nucleic acid extraction kit (MP Biomedicals, USA) following the manufacturer’s instructions. Phylogenetically diagnostic sequences were amplified using the bacterial 16S rRNA universal primers 515F (5ʹ-GTGCCAGCMGCCGCGG-3ʹ) and 907R (5ʹ-CGTCAATTCMTTTRAGTTT-3ʹ). Amplified DNA was verified by electrophoresis of PCR mixtures in 1.2 % agarose in 1X TAE buffer and purified using a protein nucleic acid detector (Bio-RAD, USA).

Sequencing and phylogenetic analysis

Samples were sent for sequencing on a Miseq platform. Sequencing data were cleaned using the software package Qiime and then clustered to operational taxonomic units (OTUs) with a complete linkage algorithm at a 97% sequence identity level. Abundance-based coverage estimators, observed_otus, the Chao1, Shannon, and Simpson parameters were estimated for alpha diversity analysis. A rarefaction curve was also analyzed using QIIME software (http://qiime.org/scripts). Taxa with proportions <0.01% were grouped as “others”. With the VEGAN package in the integrated suite of software facilities R (Oksanen et al. 2009), redundancy analysis (RDA) was used to examine the correlations between community variations, community functions and environmental parameters. The heatmap.2 program within the gplots package was used to paint the heat map. PICRUSt, a bioinformatics tool designed to address the functional potential in different sites using 16S ribosomal DNA sequences (Langille et al. 2013).

For this analysis, the closed-reference OTU picking protocol was performed using QIIME1.9.0 (Caporaso et al. 2010). Sequences are aligned with the Greengenes database (vers. 13.5) (Desantis et al. 2006). The OTU table was created after rarefying samples to 29652
sequences and the gene copies were normalized, then using the PICRUSt for further analysis. This normalization helped us avoid overestimation of some groups of microorganisms (Urbanová et al., 2014). For example, without normalization the estimate of the relative abundance of Proteobacteria can be up to twice as high. It performed functional analysis for COGs and KEGG ortholots. Here, we used the KEGG ortholots (KOs). The relative abundance of functional categories was generated using the OTU table of assigned taxa and their relative distribution in different samples (Urbanová et al., 2014). In the KEGG database, functions were grouped into three level subgroups based on different KEGG functional gene ontology affiliation (i.e. metabolism, cellular processes, environmental processing).

Accession numbers

All the sequences in this study have been submitted to the NCBI-SRA public database (http://www.ncbi.nlm.nih.gov/sra/SPR104573) under the ID: SRP104573 (all the twelve samples of the Changjiang Estuary).

RESULTS

Environmental parameters and bacterial counts

The environmental parameters measured are given in Table 1. Samples from site A0102 had the highest COD while A0502 had the lowest; B0202 had the highest NH$_4^+$ concentration while A0302, A0402 and B0402 were the lowest; DO was highest at site C0402 and lowest at C0102; NO$_2^-$ concentration was relatively lower at all sites. DAPI fluorescence direct counts of 1 or 2 mL of seawater are shown in Table 1; site A0102 had the highest total bacterial count and C0102 the lowest.

Bacterial community structure

A total of 445025 16S rRNA gene sequences were obtained from the 12 sample sites. The sequences with insufficient quality or sequences that could not be adequately assigned were
not included, such as chimera sequences. 442894 sequences were retained for further analysis.

To compare the diversities and richness of the bacterial communities, Chao1 and ACE estimates and the Shannon-Weaver diversity index were calculated (Table 2). The richness, estimated by number of OTUs, Chao1 and ACE indices, showed that the highest bacterial richness was at site A0102 with the lowest at A0502. Similarly, the Shannon and Simpson diversity indices indicated that site A0102 had the highest bacterial diversity with lowest diversity at A0202. Rarefaction curves show the diversity and richness of each sample (Figure 2). A0102 had the highest bacterial diversity and richness while A0502 was the lowest.

QIIME software was used to identify sequences obtained or to find sequences similar to those obtained here. Overall, there were 36 phyla in the samples. The composition and structure of the bacterial communities in the different samples were compared at the phylum level (Figure 3): Proteobacteria (67.8%), Acidobacteria (0.8%), Actinobacteria (12.2%), Bacteroidetes (11.7%), Chlorobi (0.1%), Chloroflexi (0.2%), Cyanobacteria (2.5%), Firmicutes (0.1%), Gemmatimonadetes (0.3%), Nitrospirae (0.2%), OP3 (0.1%), PAUC34f (0.1%), Planctomycetes (0.1%), SAR406 (2%), SBR1093 (0.1%), Tenericutes (0.1%), Verrucomicrobia (0.8%), ZB3 (0.2%), and unassigned (0.3%). “Others” were present in very low abundance. In current study, Proteobacteria were the major component in each library, followed by Actinobacteria and Bacteroidetes. Alphaproteobacteria (38.7%) and Gammaproteobacteria (22.8%) were the predominant classes observed and were in each samples. Firmicutes, SBR1093, Tenericutes, and ZB3 were also detected in all samples (with very low abundance).

**Similarity of bacterial communities in the seawater samples**

The heatmap in Figure 4 shows the abundance of the bacteria in the 12 samples. The different samples could be clustered into six groups: A0202; B0302/C0102; A0402/C0402;
A0502/B0402; A0102; A0302/B0102/B0202/C0302. The communities of B0302 and C0102 were highly similar, which had Proteobacteria and Actinobacteria as the dominant phyla. A0402 and C0402 were similar and had Proteobacteria, Actinobacteria and Cyanobacteria as the dominant phyla while B0402 and A0502 had the Proteobacteria, Bacteroidetes and SAR406 as the dominant phyla. The A0302, B0102, B0202 and C0302 were highly similar, which had the Proteobacteria, Actinobacteria and Bacteroidetes as the dominant phyla. However, the communities of samples A0102 and A0202 were distant from the other samples. A0102 had the Acidobacteria, Proteobacteria and Actinobacteria as the dominant phyla while the A0202 had the Proteobacteria and Actinobacteria as the dominant phyla. The abundance of bacteria was obtained directly from the heatmap. Proteobacteria were dominant, followed by Actinobacteria and Bacteroidetes, while the least abundant were Firmicutes, Chlorobi, PAUC34f, OP3 and SBR1093.

Contribution of environmental factors to bacterial community structure and functional genons

Our study indicated that environmental factors might be important determinants of structure and function of microbial communities in the surface water samples. To explore how environmental parameters influenced the bacterial community composition and functional diversity, RDA was performed. Result indicated that PO$_4^{3-}$ and DO showed positive relationships with taxonomic composition (Figure 5). The first two axes explained 39.4% of the taxonomic information. The envfit showed that PO$_4^{3-}$ had significant influence on the bacterial communities ($P < 0.05$) (table 3). Meanwhile, the PO$_4^{3-}$, DO and COD were significantly associated with microbial functional diversity. The first two axes explained 72% of the function information. Envfit showed that PO$_4^{3-}$ and COD had significant influences on the functional
genes (P < 0.05) (Table 4). PO$_4^{3-}$ was identified as a major environmental factor structuring the microbial community and contributing to the function of microbial communities. The mantel test (Table 5) indicated that functional genes showed higher relationships to environmental factors than taxonomic composition.

**Metagenome analysis**

Based on the 16S rRNA gene copy number of detected phylotype, we predicted the functional profiles of bacterial communities among the 12 water samples. The relative abundance of functional profiles were similar in most of samples. We observed that amino acid metabolism, carbohydrate metabolism, membrane transport and energy metabolism were pronounced enriched in all samples (Figure 6). We also analyzed functional profiles that were involved in the bacterial community adaptation to environment and nutritional conditions (Figure 7). At the individual pathway level, we found that oxidative phosphorylation, carbon fixation pathways in prokaryotes, photosynthesis, nitrogen metabolism, carbon fixation in photosynthetic organisms, polycyclic aromatic hydrocarbon degradation, sulfur metabolism and methane metabolism were enriched in all samples. However, there were also pronounced differences of the functional categories among the samples. Photosynthesis and Polycyclic aromatic hydrocarbon degradation were markedly enriched in A0102. Oxidative phosphorylation, Carbon fixation in photosynthetic organisms, Carbon fixation pathways in prokaryotes and Nitrogen metabolism were pronounced enriched in C0102. These two sites were close to the estuary.

**DISCUSSION**

**Bacterial abundance and its correlation with environmental factors**

The bacterial biomass was counted by DAPI fluorescence direct enumeration. Sample
A0102 had the highest total bacterial counts while A0402 had the lowest. Bacterioplankton are typically present at $10^8$–$10^9$ cells/L in seawater (Zeng et al. 2014). The average bacterial biomass was $1.76 \times 10^5$ cells/mL, which conformed with previous reports: Zhao (2003) found bacterial abundance ranging from $3.05 \times 10^5$ cells/mL to $1.36 \times 10^6$ cells/mL in the East China Sea (Zhao et al. 2003). However, our result was higher than the values reported by Lin (1998) from the Pacific Ocean and Prydz Bay in Antarctica (Lin & Zeng 1998). In recent years the Changjiang Estuary has become extensively polluted, which provides rich nutrients for bacterial growth. This may lead to the higher bacterial abundance in the Changjiang Estuary. We also observed that the bacterial abundance was higher in the estuary compared to the open sea. For example, sampling sites A0102, A0202 and A0302 (Figure 1) had higher bacterial biomass than A0402 and B0402. Thereby, we deduced that the estuary had higher nutrient levels than the open sea. This may result from the effect of coastal flow. The Subei coastal current and Yellow Sea Coastal Water could bring ample nutrients to the estuary. Yang et al. reported that under the effect of coastal flow on the coastal side, the content of oxygen and nutrients near the Changjiang Estuary was obviously higher than that in the open sea (Yang et al. 2008). Additionally, when the northward TWC, southward Subei coastal current, and Yellow Sea coastal current mix with the freshwater of the Changjiang River in the estuary, there could be ample nutrients to contribute to bacterial growth. Nutrients are supplied to the Changjiang Estuary by the Changjiang River, the TWC respectively, at different times, by different means and with different intensities (Yang et al. 2008). The Changjiang Estuary is a complex ecosystem, the required material sources for microbial growth is varied, including the terrigenous runoff, vertical mixing and resuspension (Shiah & Ducklow 1995). Thereby, the abundance of bacteria and its distribution may be influenced by many environmental factors.
Diversity of bacteria

We used 16S rDNA gene sequences, which were analyzed according to a taxon cutoff set at 97% similarity, to assess the biodiversity of bacterial communities in this environment. Proteobacteria, Actinobacteria and Bacteroidetes were the dominant phyla. Alphaproteobacteria and Gammaproteobacteria were the most abundant subphyla (classes), which was consistent with previous study: Feng et al. reported that Alphaproteobacteria (23.4%) and Gammaproteobacteria (31.7%) were the most abundant phylum in the Changjiang Estuary, followed by Firmicutes (6.4%), Bacteroidetes (4.6%) and Actinobacteria (4.1%) (Feng et al. 2009). Skeiguchi (2002) also found that Alphaproteobacteria and Gammaproteobacteria were the predominant taxa in the Changjiang Estuary (Sekiguchi et al. 2002). Proteobacteria was the dominant phylum in all sites around Xiamen Island (Shan et al. 2015) (including 1569 OTUs, accounting for 49.62%–76.84% of the total reads at different sites, among which Alphaproteobacteria and Gammaproteobacteria were the predominant classes in seawater).

In our study, Alphaproteobacteria was the dominant class and distributed in all samples with high abundance, which concurred with previous studies showing that Alphaproteobacteria were the dominant group in open and coastal ocean environments (Bernhard et al. 2005; FO et al. 1999). Previous study found that Alphaproteobacteria related to the process of Sulfide(Li et al. 2008). Betaproteobacteria had their highest abundance at site A0102. Previous studies have found that Betaproteobacteria had positive correlations with low salinity (Liu et al. 2009; Mosier & Francis 2008). In our study, this site was close to the offshore. The east China sea coastal current, originated in the Yangtze estuary and hangzhou bay area, the dilute water from the Yangtze river and the qiantang river, all of them flow to the south, which led to the low salinity
in the south of Fujian and Zhejiang coast. This may lead to the high richness of Betaproteobacteria in the A0102. Gammaproteobacteria were distributed in all sites with relatively high abundance. Gammaproteobacteria played important role in the anoxic nitrification-manganese reduction processes (Freitag & Prosser 2003). Hence, the high abundance of Gammaproteobacteria may be closely related to the carbon-nitrogen cycle, which was conformed with our study that the carbon fixation pathways in prokaryotes and nitrogen metabolism were significantly enriched in all samples. Deltaproteobacteria was also distributed in all samples but with very low abundance. Deltaproteobacteria can degrade naphthalene, alkylbenzenes and benzene (Musat et al. 2009). Actinobacteria and Bacteroidetes were also dominant taxa in all sites. Actinobacteria play important role in the degradation of organic pollutants (He-Yang et al. 2012), which was consistent with our study that Actinobacteria had their highest abundance at site C0302 and the nitrodoluene degradation, chloroalkane and chloroalkene degradation and naphthalene degradation were significantly enriched in C0302. Bacteroidetes was able to degrade biological macromolecules, such as, Chitin, AGAR, DNA and cellulose (Reichenbach 2006). In our study, xenobiotic biodegradation pathways were enriched in all samples including Chitin and AGAR. Cyanobacteria live in relatively pristine coastal environments and play an important role in primary production (Sun et al. 2003). Acidobacteria were mainly distributed at sampling sites A0102, A0202, B0102 and B0202. Barns et al. reported that Acidobacteria have the ability to withstand metal-contaminated, acidic environments and may also be widely distributed in radionuclide-contaminated environments (Barns et al. 2007). In addition, such bacteria are commonly abundant in the terrestrial environment, such as the soil (Wang et al., 2010). In our study, these four sites were close to the estuary and affected by terrigenous, which led to the higher abundance of Acidobacteria.
The relationship between community structure, functional diversity and the environmental factors

Bacterial communities are always closely related to their environments and factors such as salinity, temperature and nutrients can be used as an indicator of marine ecosystem status (Andersson et al. 2010; Herlemann et al. 2011). In our study, the results of RDA showed that DO and PO$_4^{3-}$ were the most important environmental factors that influenced the distribution of the bacterial community structure in surface seawater of the Changjiang Estuary and adjacent areas. Bacterial communities in the water and sediment of the Tama River were influenced by organic matter in the water (Sugita et al. 1983). Joshua et al. reported that phosphate concentration predicted the diversity of bacteria (Ladau et al. 2013). The bacterial community structure responds quickly to changes in DO (Yan et al. 2008). Consistent with these findings, PO$_4^{3-}$ and DO had the most significant correlations with the bacterial communities in our study. In our study, we found that NH$_4^+$ had no correlation with the bacterial community structure. Recent study has revealed that the bacterial ability of using NO$_3^-$ was higher than we thought (Middelburg & Nieuwenhuize 2000). The high abundance of NO$_3^-$ could meet the needs of the N source for the bacteria, which weakens the bacteria dependence on nitrogen. We also observed that PO$_4^{3-}$, DO and COD were markedly associated with microbial functional diversity. Interestingly, from the Mantel test, we found that the functional genes was significantly correlated with environmental conditions than the taxonomic genera, indicating that functional gene patterns may be more sensitive to environmental conditions than taxonomic composition. Similarly, functional genes appeared to be more appropriate than ‘species’ information in addressing questions regarding bacterial community assembly (Burke et al. 2011). Bacterioplankton are essential players in the release of phosphorus and nitrogen fixation in
We observed that PO$_4^{3-}$ was the most important factors affecting the bacterioplankton community function and composition. The Changjiang Estuary and its adjacent areas are a complex area system which mixes with the TWC. The TWC originated from the Taiwan Strait, extended along the coast of Fujian and Zhejiang Provinces to the north and met with Changjiang Estuary. Therefore we speculated that the TWC may bring abundant phosphate to the area. Similarly, the Changjiang River could also transport phosphate to the Changjiang Estuary. Yang et al. found that the sources of phosphate in the Changjiang estuary included the Changjiang River, the TWC, the cyclone-type eddy and the 32°N upwelling (Yang et al. 2008). These may result in the high phosphate concentration in the Changjiang Estuary, variation of which correlated significantly with the bacterial communities and functions.

**Predictive Functional Analysis**

PICRUSSt provides a prediction of microbiome function based on marker genes, but not an actual measurement of such function. The NSTI scores was used to evaluate the predictive accuracy of PICRUSSt. The PICRUSSt predictions of genomes tend to be less accurate in poorly environment, for there are relatively few reference genome sequences available (Cleary et al. 2015). In our study, the NSTI was relatively lower. Langille et al. reported that the accuracy of PICRUSSt decreased with increasing NSTI scores. We found that oxidative phosphorylation, carbon fixation pathways in prokaryotes, photosynthesis, nitrogen metabolism, carbon fixation in photosynthetic organisms, polycyclic aromatic hydrocarbon degradation, sulfur metabolism and methane metabolism were enriched in all samples. The polycyclic aromatic hydrocarbon degradation was significantly enriched in A0102 and C0102. However, it was lower in B0402 and A0502. Previous study reported that petroleum content decreased from the alongshore to the
open sea (Lei et al. 2014), which was consistent with our study. Site A0102 and C0102 were close to the shore while B0402 and A0502 were in the open sea. The nitrogen metabolism was markedly enriched in all samples. The relationship between taxonomic composition and nitrogen sources was reported in some other river system (Clara Ruiz-González et al. 2013). This correlation may be attributed to the Proteobacteria, which includes many members involved in nitrogen cycling (Breitbart et al. 2009; Yang et al. 2013). However, further studies are needed to determine how the detected bacterial members stimulate nutrient cycling processes in this ecosystem (Yan et al. 2015). The photosynthesis was markedly enriched in A0102. Cyanobacteria is widely distributed in the ocean and can do photosynthesis (Sun et al. 2003). In our study, we observed that Cyanobacteria had the highest abundance in A0102, which may lead to the enrichment of photosynthesis. The phosphonate and phosphinate metabolism was also enriched in all samples. With the development of science and technology, eutrophication phenomenon was increasingly serious in the Changjiang Estuary. It may lead to the enrichment of the phosphonate and phosphinate metabolism.

CONCLUSIONS

The bacterial biomass was high in surface seawater of the Changjiang Estuary and surrounding areas and differed among sites. We detected 36 phyla, including Proteobacteria, Acidobacteria, Actinobacteria, Bacteroidetes and unclassified, as well as others with low abundance. Proteobacteria was the dominant phylum, followed by Actinobacteria and Bacteroidetes. Diversity analysis indicated that sampling site A0102 had the highest diversity of bacteria. RDA result showed that PO$_4^{3-}$ was the main environmental factors that influenced the distribution of bacterial communities and functions. Analysis of predictive metagenomic showed that Oxidative phosphorylation, Carbon fixation pathways in prokaryotes, Photosynthesis, Nitrogen...
metabolism, Carbon fixation in photosynthetic organisms, Polycyclic aromatic hydrocarbon degradation, Sulfur metabolism, Methane metabolism were enriched in all samples. These findings expand our current understanding on bacterial structure and function in the Changjiang Estuary and its adjacent areas.

Acknowledgements

This work was supported by the National Natural Science Foundation of China [grant numbers 31270160, J1310037] and the Natural Science Foundation of Zhejiang Province, China [grant number LY12C03003].

References


Feng BW, Li XR, Wang JH, Hu ZY, Meng H, Xiang LY and Quan ZX. 2009. Bacterial diversity of water and sediment in
the Changjiang estuary and coastal area of the East China Sea. *FEMS microbiology ecology* 70: 80-92.


Shan D, Wei G, Li M, Wang W, Li X, Gao Z and Shao Z. 2015. Distribution and diversity of bacterioplankton...
Table 1. Environmental parameters, bacteria direct counts, and location of the 12 sample sites

<table>
<thead>
<tr>
<th>Sample</th>
<th>NH$_4^+$ (mg/L)</th>
<th>PO$_4^{3-}$ (mg/L)</th>
<th>COD (mg/L)</th>
<th>DO (ml/L)</th>
<th>NO$_2^-$ (mg/L)</th>
<th>Total bacterial count (cells/L)</th>
<th>Longitude</th>
<th>Latitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0102</td>
<td>0.07</td>
<td>0.05</td>
<td>2.86</td>
<td>5.48</td>
<td>0</td>
<td>3.40×10$^5$</td>
<td>122°12′42″</td>
<td>30°58′48″</td>
</tr>
<tr>
<td>A0202</td>
<td>0.11</td>
<td>0.05</td>
<td>3.63</td>
<td>5.66</td>
<td>0.01</td>
<td>2.62×10$^5$</td>
<td>122°20′42&quot;</td>
<td>30°56′48″</td>
</tr>
<tr>
<td>A0302</td>
<td>0.03</td>
<td>0.02</td>
<td>0.8</td>
<td>5.8</td>
<td>0.01</td>
<td>1.76×10$^5$</td>
<td>122°41′18″</td>
<td>30°59′54″</td>
</tr>
<tr>
<td>A0402</td>
<td>0.03</td>
<td>0.01</td>
<td>1.37</td>
<td>6.06</td>
<td>0.01</td>
<td>1.52×10$^4$</td>
<td>122°49′30″</td>
<td>30°59′30″</td>
</tr>
<tr>
<td>A0502</td>
<td>0.04</td>
<td>0</td>
<td>0.36</td>
<td>6.46</td>
<td>0</td>
<td>2.12×10$^5$</td>
<td>122°59′54″</td>
<td>30°59′42″</td>
</tr>
<tr>
<td>B0102</td>
<td>0.12</td>
<td>0.05</td>
<td>1.56</td>
<td>4.7</td>
<td>0.01</td>
<td>1.32×10$^5$</td>
<td>122°07′00″</td>
<td>30°45′24″</td>
</tr>
<tr>
<td>B0202</td>
<td>0.52</td>
<td>0.04</td>
<td>2.02</td>
<td>5.35</td>
<td>0.01</td>
<td>2.80×10$^5$</td>
<td>122°20′30″</td>
<td>30°46′54″</td>
</tr>
<tr>
<td>B0302</td>
<td>0.11</td>
<td>0.03</td>
<td>0.92</td>
<td>5.95</td>
<td>0.01</td>
<td>2.16×10$^5$</td>
<td>122°41′48″</td>
<td>30°44′06″</td>
</tr>
</tbody>
</table>

444 communities in subtropical seawater around Xiamen Island, China. *Microbiological Research* 175: 16-23.
445 Shiah FK and Ducklow HW. 1995. Regulation of bacterial abundance and production by substrate supply and
449 Sun S, Xiao T and Yue H. 2003. Distribution character of Synechococcus spp.in the east china sea and the yellow sea
451 Urbanová Z and Bárta J. 2014. Microbial community composition and in silico predicted metabolic potential reflect
452 biogeochemical gradients between distinct peatland types. *Fems Microbiology Ecology* 90:633-646.
454 Relationships between bacterial diversity and environmental variables in a tropical marine environment, Rio de
458 Wang Y and Chen F. 2008. Decomposition and phosphorus release from four different size fractions of Microcystis
461 Reservoir Region revealed by PCR-DGGE and its relationships with environmental factors. *Journal of Environmental
462 Sciences* 20: 732-738.
464 Dam on microbial structure and potential function. *Scientific reports* 5: 8605.
465 Yang, Ping, Zhimei, Qing, Wang, Huihui, Hong and Shen. 2013. Decrease of NH$_4^+$-N by bacterioplankton
466 accelerated the removal of cyanobacterial blooms in aerated aquatic ecosystem. *Journal of Environmental Sciences*
467 25: 2223-2228.
470 Zeng YX, Yu Y, Qiao ZY, Jin HY and Li HR. 2014. Diversity of bacterioplankton in coastal seawaters of Filides
471 Peninsula, King George Island, Antarctica. *Archives of Microbiology* 196: 137-147.
474 Zhao SJ, Tian X and Yue HD. 2003. THE DISTRIBUTION OF HETEROTROPHIC BACTERIA AND SOME RELATED
475 FACTORS IN THE EAST CHINA AND YELLOW SEAS DURING FALL. *Oceanologia Et Limnologia Sinica* 34:305-311.
<table>
<thead>
<tr>
<th>Sample</th>
<th>No. of sequences</th>
<th>OTUs</th>
<th>Bias-corrected Chao1</th>
<th>Coverage Estimator</th>
<th>Shannon index(H)</th>
<th>Simpson index(D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0102</td>
<td>39176</td>
<td>1096</td>
<td>1199</td>
<td>1174</td>
<td>7.4</td>
<td>0.02</td>
</tr>
<tr>
<td>A0202</td>
<td>30107</td>
<td>482</td>
<td>549</td>
<td>516</td>
<td>4.92</td>
<td>0.16</td>
</tr>
<tr>
<td>A0302</td>
<td>40387</td>
<td>607</td>
<td>733</td>
<td>757</td>
<td>5.75</td>
<td>0.05</td>
</tr>
<tr>
<td>A0402</td>
<td>39799</td>
<td>473</td>
<td>602</td>
<td>592</td>
<td>5.63</td>
<td>0.05</td>
</tr>
<tr>
<td>A0502</td>
<td>31670</td>
<td>408</td>
<td>551</td>
<td>532</td>
<td>5.63</td>
<td>0.04</td>
</tr>
<tr>
<td>B0102</td>
<td>43482</td>
<td>749</td>
<td>851</td>
<td>850</td>
<td>6.42</td>
<td>0.03</td>
</tr>
<tr>
<td>B0202</td>
<td>36889</td>
<td>794</td>
<td>990</td>
<td>982</td>
<td>6.49</td>
<td>0.03</td>
</tr>
<tr>
<td>B0302</td>
<td>41777</td>
<td>707</td>
<td>1001</td>
<td>969</td>
<td>5.78</td>
<td>0.05</td>
</tr>
<tr>
<td>B0402</td>
<td>30993</td>
<td>490</td>
<td>615</td>
<td>599</td>
<td>5.82</td>
<td>0.05</td>
</tr>
<tr>
<td>C0102</td>
<td>33672</td>
<td>681</td>
<td>780</td>
<td>783</td>
<td>6.32</td>
<td>0.03</td>
</tr>
<tr>
<td>C0302</td>
<td>32430</td>
<td>638</td>
<td>827</td>
<td>791</td>
<td>5.9</td>
<td>0.04</td>
</tr>
<tr>
<td>C0402</td>
<td>44643</td>
<td>532</td>
<td>674</td>
<td>689</td>
<td>5.61</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 3. Redundancy analysis of environmental factors and bacterial community structure in the Changjian estuary and adjacent areas.

|          | RDA1          | RDA2          | r²           | Pr(>|r|) |
|----------|---------------|---------------|--------------|---------|
| NH₄⁺     | -78.33×10⁻²   | 62.12×10⁻²    | 32.99×10⁻²   | 0.171   |
| PO₄³⁻    | -99.99×10⁻²   | 1.27×10⁻²     | 49.06×10⁻²   | 0.049*  |
| COD      | -95.65×10⁻²   | 29.18×10⁻²    | 26.40×10⁻²   | 0.24    |
| DO       | 79.55×10⁻²    | 60.60×10⁻²    | 47.27×10⁻²   | 0.054   |
| NO₃⁻     | 5.32×10⁻²     | 99.86×10⁻²    | 38.92×10⁻²   | 0.106   |

*P<0.05. Number of permutations: 999.

Table 4. Redundancy analysis of environmental factors and functional genes in the Changjiang estuary and adjacent areas.
Table 5. Summary results from Mantel tests performed at the functional gene level or phylogenetic genus level.

|                  | RDA1        | RDA2        | $r^2$       | Pr(>|r|) |
|------------------|-------------|-------------|-------------|----------|
| NH$_4^+$         | 73.76×10$^{-2}$ | -67.52×10$^{-2}$ | 11.06×10$^{-2}$ | 0.467    |
| PO$_4^{3-}$      | -5.82×10$^{-2}$ | -99.83×10$^{-2}$ | 49.40×10$^{-2}$ | 0.037*   |
| COD              | -74.15×10$^{-2}$ | -67.09×10$^{-2}$ | 52.16×10$^{-2}$ | 0.036*   |
| DO               | -53.45×10$^{-2}$ | 84.52×10$^{-2}$   | 53.17×10$^{-2}$ | 0.082    |
| NO$_2^-$         | -57.48×10$^{-2}$ | -81.83×10$^{-2}$ | 15.20×10$^{-2}$ | 0.485    |

*P<0.05. Number of permutations: 999.
Figure 1. Distribution of the 12 sampling sites in the Changjiang Estuary and adjacent areas.
Figure 2. Rarefaction curves for the 12 samples
Figure 3. Community structure of the 12 water samples at the phylum level
Figure 4. Heat map of the 12 water samples based on the abundance similarity of the bacteria. Columns represent the different samples; rows represent the bacteria. Red represents the highest bacterial abundance, light blue the lowest. The "*" represent the relative abundance of samples > 0.1 while the "+" represent the relative abundance > 0.0.
Figure 5. Redundancy analysis (RDA) shows the relationships between environmental variables and the bacteria communities (a); relationships between environmental variables and the functional genes (b).
Figure 6. Heatmap showing the differences among the 12 investigated communities based on the KEGG orthology groups. Columns represent the different samples; rows represent the bacteria. Red represents the highest bacterial abundance, light blue the lowest. The “*” represent the gene counts > 600000 while the “+” represent the gene counts > 300000.
Figure 7. The relative abundance of some imputed functional profiles in the 12 water samples.