Comparison of the community structure of planktonic bacteria in ballast water from entry ships and local sea water in Xiamen Port

Ying Ma a, Hejian Xiong b, Senming Tang c,*, Qingshuang Yang d, Minjuan Li a

a Fisheries College of Jimei University, Xiamen 361021, China
b Bioengineering College of Jimei University, Xiamen 361021, China
c The Third Institute of Oceanography, SOA, Xiamen 361005, China
d Xiamen Entry–Exit Inspection and Quarantine Bureau, Xiamen 361012, China

Received 23 July 2008; received in revised form 11 August 2008; accepted 14 September 2008

Abstract

In this study, the bacterial community structures in samples of ballast water collected from a ship from Singapore and of local sea water collected from Xiamen Port were compared using restriction fragment length polymorphism (RFLP) and 16S rDNA sequence analysis. Except for dominant α-Proteobacteria that are common to both systems, the bacterial community structures of the two systems were quite different. Most of the clones derived from the different systems were grouped into different phylogenetic clusters, and the systems share only one common RFLP pattern. The ballast water, which is likely from clean offshore waters, contains sequences specific to α- and γ-Proteobacteria. Phylogenetic analysis revealed that the ballast water contained sequences belonging to attached bacteria and bacteria commonly found in the open sea, as well as many novel sequences. In addition, no known pathogenic bacteria were detected in the ballast water samples. Conversely, water samples from Xiamen Port were apparently affected by the near shore environments. Specifically, in addition to α- and γ-Proteobacteria, water from Xiamen Port contained β- and δ-Proteobacteria, Synechococcus, Bacteroidetes and Actinobacteria, which are common in coastal environments. Additionally, four pathogenic bacterial sequences and one plasmid sequence of a potential red tide forming alga were detected in the water from Xiamen Port, which suggests that the local sea water is polluted. The results of this study can be used as background information to assess the risk associated with the introduction of non-indigenous species to local systems and to establish ballast water management systems.

© 2009 National Natural Science Foundation of China and Chinese Academy of Sciences. Published by Elsevier Limited and Science in China Press. All rights reserved.

Keywords: Xiamen; Ballast water; Bacteria; 16S rDNA; Restriction fragment length polymorphism (RFLP)

1. Introduction

It is estimated that 2–3 billion tons of ballast water are carried around the world each year [1]. As a result, translocation of organisms by ships is one of the most important issues threatening the naturally evolved biodiversity of local aquatic systems, and the consequences of the unintentional introduction of alien species associated with ships have gained increased attention in recent years. Ballast water contains a diverse mix of phytoplankton, including HAB, cysts and eggs of zooplankton, benthic organisms, and bacteria and viruses [2]. Although many studies that have been conducted to evaluate ballast invaders have focused on macro-organisms such as metazoans [3], micro-organisms present in ballast water have largely been ignored. This lack of attention is likely because microbial invaders cannot be detected without the aid of compound microscopes and their presence is generally only noticed when they have large impacts, such as during red tides or outbreaks of illness [4]. Despite this lack of attention, microbes are the most abundant organisms in the environment, arrive in the
greatest numbers in ballast tanks, and probably have the greatest chances of survival upon introduction into a new environment. Ruiz et al. [5] reported that the mean number of bacteria and virus-like particles in the ballast water of vessels arriving in Chesapeake Bay from foreign ports was $8.3 \times 10^6$ per liter and $7.4 \times 10^8$ per liter, respectively. However, to date, most studies conducted to evaluate micro-organisms in ballast water have primarily focused on harmful alga [6,7], pathogens and viruses [5,8,9], and there is a lack of information regarding other non-indigenous micro-organisms. As a result, it is difficult to fully assess the ecological and economic impacts of invading species on local communities.

Xiamen is an important shipping center on the southeastern coast of China. It has been reported that Xiamen Port receives more than 4 million tons of ballast water of foreign origin annually, excluding ballast water received from other domestic ports [10]. Some of this ballast water is taken on in areas that are known to have been impacted by infectious diseases, which may result in the dispersal of pathogens or harmful algae. In this study, we collected a ballast water sample from a ship arriving in Xiamen Harbor and then evaluated the bacterial genetic diversity using a 16S rRNA gene clone library and sequences analysis. We then used this information to compare the bacterial community structure of the ballast water to that of the local seawater. It is hoped that the information generated in this study can be used as background information to enable better assessment and management of the introduction of non-indigenous species, including potentially pathogenic micro-organisms, via ballast-water transfer.

2. Materials and methods

2.1. Sample collection

A ballast water sample was collected from a ship anchored in Xiamen Harbor, China (24°29’N, 118°04’E), in September of 2007. The source region of the ballast water was Singapore. The ballast water sample comprised subsamples collected from the subsurface at five different locations onboard the ship. Five liters of the water were then filtered through 47 mm diameter polysulfone filters with a pore size of 0.2 μm (PALL, Ann Arbor, USA) at <0.03 MP. Subsurface seawater (5 l) from Xiamen Harbor was also filtered using the same procedure. The filters were then immediately frozen and stored at −20°C until DNA extraction.

2.2. DNA extraction, PCR amplification and clone library construction

The filters containing the samples were cut into small pieces and then incubated in 1.5 ml of lysis solution (45 mM glucose, 23 mM Tris [pH 8.0], 59 mM EDTA) containing 0.5 mg/ml lysozyme. After being incubated at 37°C for 2 h, the lysates were transferred to another sterile centrifuge tube, after which the filter species were rinsed again with 0.5 ml of lysis buffer. Next, the lysates were pooled and the cells were then further lysed by the addition of 1/100 volume of proteinase K (0.5 mg/ml) and 1/10 volume of sodium dodecyl sulfate (1%, SDS) followed by incubation for 30 min at 55°C. The DNA was extracted from cell lysates by the phenol–chloroform extraction method. The PCR was performed to amplify the bacterial 16S rDNA. Primer sequences were 27F (5’-AGAGTTTGATCCTGGCTCAG-3’) and 1492R (5’-GGTACCCTTGTAGACTT-3’). The amplification reaction mixture consisted of 0.5 μM of each primer, 200 μM dNTPs, 5 μl of 10× PCR buffer, 1 unit of Taq DNA polymerase (TaKaRa Biotechnology Co., Dalian, China) and 2 μl of DNA template. The amplification conditions were as follows: initial denaturation at 94°C for 5 min, followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min, with a final extension at 72°C for 10 min. The amplified products were then gel-purified and ligated into the pMD18-T vector (TaKaRa Biotechnology Co., Dalian, China), after which they were transformed into competent Escherichia coli DH5α cells. The ampicillin-resistant clones were then randomly picked and screened for inserts by performing colony PCR using M13 primers specific for the vector (Invitrogen, Shanghai, China).

2.3. PCR-RFLP screening and statistical analysis of the clone libraries

A total of 65 clones (35 from water collected from Xiamen Port and 30 from ballast water) with inserts of the expected length were selected for subsequent RFLP analysis. A 1:10 dilution of the colony PCR products that had been amplified using the M13 primer was then re-amplified with the 27F and 1492R primers under the same PCR conditions described above. The PCR products were then separately digested using the restriction endonucleases, HhaI and AfaI (TaKaRa Co., Dalian, China). The restriction fragment patterns were then visualized on a 2.0% agarose gel, after which the clones were discriminated according to their RFLP patterns. The phylotype diversity in the clone libraries obtained through PCR-RFLP analysis was then subjected to statistical analysis. Specifically, the following indices were calculated: (1) Taxa, the total number of RFLP patterns in each library; (2) individuals, the total number of clones examined; (3) coverage [11], which was determined using the following equation: Coverage = 1 − (N/Individuals), where N is the number of clones that occurred only once; (4) diversity indices (dominance, evenness, Shannon), which were calculated using the statistical program, PAST (http://folk.uio.no/ohammer/past).

2.4. DNA sequencing and phylogenetic analysis

Representative clones showing unique RFLP patterns were selected for sequencing, which was conducted using an ABI model 377 automated DNA sequence analyzer.
(Applied Biosystems, Perkin-Elmer) and the 27F sequencing primer. All nucleotide sequences were checked for putative chimeras using RDP CHIMERA_CHECK [12] and then compared with known 16S rDNA sequences in the database using a BLASTN search (http://www.ncbi.nlm.nih.gov/BLAST/). Multiple alignments were then performed using the neighbor-joining algorithm in the ClustalX software [13], after which a phylogenetic tree was constructed using the MEGA3 software [14]. Bootstrap values were obtained with 100 replicates.

### 2.5. Nucleotide sequence accession numbers

The cloned sequences have been deposited in GenBank under Accession Nos. EU877627 to EU877670.

### 3. Results

#### 3.1. PCR screening, pattern frequency and statistical analysis of the two bacterial 16S rDNA clone libraries

After screening the colonies for inserts by colony PCR using the M13 primer, a total of 65 clones were subjected to RFLP analysis, 35 of which were from the Xiamen Port water library and 30 of which were from the ballast water library (Table 1). Restriction analysis indicated that the Xiamen clone library contained 25 RFLP patterns, and that the ballast water library contained 19 patterns. Although the number of clones used for RFLP analysis from the ballast water (30 clones) was smaller than the number of clones from Xiamen Port (35 clones), the coverage value for the ballast water (53%) was higher than the coverage value calculated for water collected from Xiamen Port (34%) (Table 1). This result indicates that the sampling size of the ballast water was more adequate than that of Xiamen Port water, and further sampling of Xiamen Port would reveal more unique clones. However, the coverage values of both clone libraries were not high, indicating that the diversity was likely underestimated in this study.

Details regarding the frequency of the RFLP patterns for different clones are shown in Fig. 1. The most abundant RFLP pattern in water from Xiamen Port corresponded to nine clones (25.1% of the total clones), whereas the most abundant RFLP pattern in the ballast water only corresponded to six clones (20% of the total clones). Furthermore, there were 23 unique RFLP patterns in the libraries corresponding to water from Xiamen Port (92% of the total patterns and 65.7% of the total clones), whereas there were 14 unique RFLP patterns (73.7% of the total patterns and 46.7% of the total clones) observed in the ballast water (Fig. 1). These results are in accordance with the higher dominance and lower evenness values that were observed for water from Xiamen Port (Table 1). In addition, greater than 70% of the RFLP patterns in each clone library were observed only one time, which reflects the significant genetic diversity of the bacterial populations in both water systems. Additionally, the Shannon index was higher for water from Xiamen Port than for the ballast water (Table 1), suggesting that the water in Xiamen Port has a higher overall diversity than the ballast water. These results are supported by the figure describing the RFLP frequency (Fig. 1). Specifically, the trend line generated for the ballast water sample is smoother, having a lower starting value and being shorter than the trend line for water from Xiamen Port. These findings suggest that the concentration of the most dominant bacteria in the ballast water was lower than that of the most dominant bacteria in the water from Xiamen Port. Furthermore, these results indicate that the bacterial community structure of the ballast water had a higher evenness and lower diversity than that of water from Xiamen Port. In addition, the two clone libraries shared only one common RFLP pattern (Fig. 1, 3\(^a\) and 7\(^a\)), and this RFLP pattern only accounted for 4% and 5% of the total RFLP patterns observed in water from Xiamen Port and ballast water, respectively. Taken together, these findings indicate that the bacterial groups in the two clone libraries differed significantly.

#### 3.2. Comparison of bacterial community structures in the two clone libraries

One or two clones with unique RFLP patterns were selected from each library and were subjected to sequencing. After removing the chimeric sequences and a few undetermined sequences, a total of 28 and 26 sequences were obtained from the Xiamen Port water and ballast water clone libraries, respectively. Sequence analysis revealed that clones within the same RFLP pattern shared \(>97\%\) sequence similarity, and that clones with different RFLP patterns had \(<97\%\) similarity, suggesting that the division of the RFLP patterns was reliable. Sequences were then assigned to major groups based on BLAST similarities and phylogenetic analysis. The species-composite clone libraries are shown in Fig. 2.

All the cloned sequences fell into seven major lineages of the bacterial domains, the \(\alpha\)-, \(\beta\)-, \(\gamma\)-, and \(\delta\)-Proteobacteria, Bacteroidetes, Actinobacteria, Cyanobacteria (Synechococcus), as well as a group originating from chloroplasts of the eukaryote Skeletonema pseudocostatum (Fig. 2). The percentage of 16S rDNA sequences from each group in the total rDNA pool indicated that the community structures

<table>
<thead>
<tr>
<th>Clone library</th>
<th>Individuals</th>
<th>Taxa</th>
<th>Coverage (%)</th>
<th>Dominance</th>
<th>Evenness</th>
<th>Shannon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ballast water</td>
<td>30</td>
<td>19</td>
<td>53</td>
<td>0.08444</td>
<td>0.8076</td>
<td>2.732</td>
</tr>
<tr>
<td>Xiamen Port water</td>
<td>35</td>
<td>25</td>
<td>34</td>
<td>0.09224</td>
<td>0.7242</td>
<td>2.896</td>
</tr>
</tbody>
</table>

Table 1
Statistical analysis of the bacterial 16S rDNA clone libraries of ballast water and water from Xiamen Port.
of the two clone libraries were different. Specifically, clones corresponding to α- and γ-Proteobacteria were found in both libraries, while clones corresponding to the remainder of the bacterial groups identified were only observed in the Xiamen clone library. Furthermore, the two libraries shared only one common RFLP pattern, which indicates that the species components within the α- and γ-Proteobacteria differed between the two clone libraries (Fig. 1). This finding was supported by the results of the phylogenetic analysis, which revealed that most of the clones derived from different libraries were grouped into different clusters (Fig. 3). When the clone abundance of the different groups was evaluated, even though α-Proteobacteria were dominant in both libraries, it accounted for only 48% of the clones in the Xiamen Port library, while it accounted for greater than 60% of the clones in the ballast water clone library (Fig. 2). Although γ-Proteobacteria were the second most dominant group in both libraries, it accounted for only 19% of the clones in the Xiamen Port library and 38% of the clones in the ballast water library, respectively.

3.3. Phylogenetic analysis of the bacterial 16S rDNA sequences in the two clone libraries

Forty-four sequences (one representative sequence from each RFLP pattern) were used for phylogenetic analysis. For the analysis, a phylogenetic tree was constructed using the neighbor-joining method. The main branches of the tree possessed high bootstrap values, indicating that the tree topology had a high confidence level (Fig. 3).

The α-Proteobacteria group primarily comprised species affiliated with Rhodobacteraceae and a few unidentified α-Proteobacterial clones. In addition, with the exception of cloned sequences XM-30 and XM-40, the α-Proteobacterial sequences in water from Xiamen Port were >97% homologous with sequences already present in Genbank. These sequences were primarily related to sequences retrieved from coastal environments including clones PL_4a9f [15], CB01D03, SIMO-4275 and B11 from coastal water (GenBank description), clone DS158 from a mangrove ecosystem, clone TH1-60 from a freshwater lake [16], and Roseovarius crassostreae CV919-312 from diseased juvenile oysters (GenBank description). These results suggest that α-Proteobacteria in the water in Xiamen Port were common to coastal environments. Conversely, more than half of the α-Proteobacteria affiliated sequences identified in the ballast water (13/19) had no close matches (<97% similarity) in GenBank, which suggests that the ballast water contained a high level of novel α-Proteobacteria. Clones that had close matches were primarily related to bacteria that were affiliated with specific hosts, bacteria from the open sea, and bacteria that had a special function. These closest relatives included bacterium DG1297, which
was attached to a dinoflagellate *Scrippsiella* sp., *Roseobacter* sp. AC2-A2, which was attached to a marine sponge, and epibiotic bacterium 11ANG521, which was attached to a squid. In addition, isolates from the ballast water were also similar to clones 6C233139 and 20162U71 from the Pacific and Atlantic oceans, respectively (GenBank description), and bacterium C49, which is known to have the ability to degrade hydrocarbons [17].

![Neighbor-joining phylogenetic tree generated based on the alignment of 16S rDNA sequences from ballast water and water from Xiamen Port. Clones from this study are indicated in boldface and are designated as BW-n and XM-n, in which BW and XM indicate that the sequences were derived from ballast water and water from Xiamen Port, respectively, and n represents the number of different clones. The remaining sequences were obtained from GenBank. Numbers in parentheses that follow the accession numbers indicate the occurrence frequency of the RFLP pattern in the clone library. Bootstrap values above 50 (100 iterations) are shown at each node. The scale bar represents the nucleotide substitution percentage.](image_url)
The γ-Proteobacteria affiliated clones were the second most dominant in both libraries, accounting for 19% and 38% of the clones in the Xiamen Port and the ballast water libraries, respectively. Three of the cloned sequences (XM-34, XM-18 and XM-19) from Xiamen Port were most closely related (98–99% sequence similarity) to sequences of the following pathogenic bacteria that are associated with disease in flounder (based on the description in GenBank): *Actinobacter junii* [18], *Vibrio pelagius* [19] and *Vibrio olivaceus*. Other clones (XM-1, XM-2 and XM-23) from water in Xiamen Port were primarily related to clones recovered from near shore waters. The dominant γ-Proteobacteria group from the ballast water library contained six clones (represented by clone BW-14). These clones were most closely related to a halophilic marine bacterium, *HAL40b*, which was isolated from the boreal sponge, *Haliclona* sp. [20]. The rest of the clones occurred only once and were closely related to the algidical bacterium, *Alteromonas* sp. A14, to *Thalassomonas ganghwensis*, which was isolated from tidal flat sediments [21], and to strain 14III/A01/015, which is isolated from the sea surface of coastal ecosystems [22].

All the remaining clones, which included β- and δ-Proteobacteria, Bacteroidetes, Actinobacteria and Cyanobacteria (*Synechococcus*), were only found in water from Xiamen Port. The β-Proteobacterial group contained three clones that were most closely related to clone PIB-41 from a freshwater lake in Austria and clone TLC-PA3-8, which was isolated from sea water in Victoria Harbor, Hong Kong [23]. In addition, three clones (represented by clone XM-36) were closely related to *Synechococcus* sp. CB11C04, which was isolated from surface water in the Chesapeake Bay (GenBank description). Additionally, two clones belonged to the Actinobacteria group, while one clone belonged to the δ-Proteobacteria group and one clone belonged to the Bacteroidetes group. These clones were all closely related to clones isolated from shallow submarine hydrothermal systems and marine sediments [24,25]. Finally, one sequence that was similar to the sequence encoding the plasmid of eukaryotic *S. pseudocostatum* (similar to the common red tide causative alga, *Skeletonema costatum*) was also detected, which indicates that the water in Xiamen Port may contain harmful algae.

4. Discussion

We utilized a clone library approach to analyze the bacterial community structure in the ballast water of a vessel arriving in Xiamen Port and then compared it with the bacterial composition of water collected from Xiamen Port. It is important to note that this technique may be influenced by possible biases introduced by PCR, such as the formation of chimeras, template annealing, and differences in the number of copies of 16S rDNA. In addition, only 65 clones were subjected to RFLP analysis, so it is unlikely that the inspected clones fully reflect the bacterial community structure of the sampled water systems. Despite these disadvantages, the clone library approach provided valuable information that has allowed us to identify the dominant bacterial composition in the ballast water, and to compare the community structure of the ballast water to that of water from Xiamen Port.

The community structure of bacteria in the water from Xiamen Port was apparently influenced by input from land-based sources of fresh water. In addition to the prevalence of α- and γ-Proteobacteria in seawater, β-Proteobacteria, which are typical in freshwater, were also found in water from Xiamen Port. The inshore characteristics of the microbial community in water collected from Xiamen Port was also supported by the phylogenetic analysis, which revealed that the cloned sequences were primarily related to sequences in GenBank that had been isolated from various coastal systems. Additionally, clones closely related to pathogens including *R. crassostreae, V. pelagius, V. olivaceus* and *A. junii*, as well as a plasmid sequence of the potential red tide causative alga, *S. pseudocostatum* were also found in the water from Xiamen Port. However, these results are not surprising because Xiamen Port shows a high degree of eutrophication and coastal contamination as a result of discharge from aquaculture and industrial and domestic sewage. Furthermore, water in Xiamen Port also contained clones related to δ-Proteobacteria, Bacteroidetes, Actinobacteria and *Synechococcus*, while only α- and γ-Proteobacteria were present in ballast water. The inshore area of Xiamen Port is a much more complex habitat than the ballast tanks of a ship. In addition, oxygen is probably never limited in water in Xiamen Port because it is constantly moving. Taken together, these factors likely enable the water in Xiamen Port to support a wider range of bacteria than ballast water.

When compared with the water in Xiamen Port, the ballast water appeared to be cleaner and contained no known pathogens. In addition, phylogenetic analysis revealed that a few sequences in the ballast water were closely related to the sequences of attached bacteria and sequences of bacteria from the open sea. These findings suggest that the ballast water was probably collected from clean offshore water, and that it is capable of supporting the growth of attached bacteria. In addition, this ballast water had been partially replaced with water from the South China Sea when the ship passed through the region in accordance with the local ballast water discharge regulations. However, even though the ballast water was clean and contained α- and γ-Proteobacteria, which already existed in the local sea water, phylogenetic analysis revealed that the species composition varied greatly between the two clone libraries. In addition, the ballast water contained several novel sequences; therefore, the potential ecological effects of the discharge of the ballast water into the local ecosystem needs further study. Moreover, one cloned sequence from the ballast water (clone BW-23) was most similar (98% sequence similarity) to a sequence obtained from the dinoflagellate, *Scrippsiella* sp., which suggests that the ballast water could also contain harmful algae.
In the present study, only one ballast water sample was analyzed, and the residual sediments within the tank were not investigated. Based on the different origins of the ballast water, future investigations that consider additional ballast water samples, as well as residual sediments in ballast tanks should be conducted to ensure that non-indigenous species are not transferred to local aquatic systems via ballast water.

Acknowledgements

This work was supported by National Special Project 908-02-03-09 and the State Key Laboratory of Marine Environmental Science (Xiamen University) Project MEL0701.

References