Accepted Manuscript

Taxonomic profiles in metagenomic analyses of free-living microbial communities in the Ofunato Bay

Md. Shaheed Reza, Atsushi Kobiyama, Yuichiro Yamada, Yuri Ikeda, Daisuke Ikeda, Nanami Mizusawa, Kazuho Ikeo, Shigeru Sato, Takehiko Ogata, Mitsuru Jimbo, Toshiaki Kudo, Shinnosuke Kaga, Shiho Watanabe, Kimiaki Naiki, Yoshimasa Kaga, Katsuhiko Mineta, Vladimir Bajic, Takashi Gojobori, Shugo Watabe

PII: S0378-1119(18)30457-8
DOI: doi:[10.1016/j.gene.2018.04.075](https://doi.org/10.1016/j.gene.2018.04.075)
Reference: GENE 42799

To appear in: *Gene*

Received date: 25 April 2018

Accepted date: 25 April 2018

Please cite this article as: Md. Shaheed Reza, Atsushi Kobiyama, Yuichiro Yamada, Yuri Ikeda, Daisuke Ikeda, Nanami Mizusawa, Kazuho Ikeo, Shigeru Sato, Takehiko Ogata, Mitsuru Jimbo, Toshiaki Kudo, Shinnosuke Kaga, Shiho Watanabe, Kimiaki Naiki, Yoshimasa Kaga, Katsuhiko Mineta, Vladimir Bajic, Takashi Gojobori, Shugo Watabe , Taxonomic profiles in metagenomic analyses of free-living microbial communities in the Ofunato Bay. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. *Gene*(2017), doi:[10.1016/j.gene.2018.04.075](https://doi.org/10.1016/j.gene.2018.04.075)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Taxonomic profiles in metagenomic analyses of free-living microbial communities in the Ofunato Bay

Md. Shaheed Reza^{a,1}, Atsushi Kobiyama^a, Yuichiro Yamada^a, Yuri Ikeda^a, Daisuke Ikeda^a,
Nanami Mizusawa^a, Kazuho Ikeo^a, Shigeru Sato^a, Takehiko Ogata^a, Mitsuru Jimbo^a, Toshiaki
Kudo^a, Shinnosuke Kaga^{b,2}, Shiho Watanabe^b, Kimiaki Naiki^b, Yoshimasa Kaga^b, Katsuhiko
Mineta^c, Vladimir Bajic^c, Takashi Gojobori^{c*}, Shugo Watabe^{a,*}

^a*Kitasato University School of Marine Biosciences, Minami-ku, Sagamihara, Kanagawa 252-0373, Japan*

^b*Iwate Fisheries Technology Center, Kamaishi, Iwate 026-0001, Japan*

^c*King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center, Thuwal 23955-6900, Kingdom of Saudi Arabia*

*Corresponding authors:

Professor Shugo Watabe, Department of Marine Biochemistry, Kitasato University,
School of Marine Biosciences, Minami-ku, Sagamihara, Kanagawa 252-0373, Japan.

E-mail: swatabe@kitasato-u.ac.jp

Professor Takashi Gojobori, Computational Bioscience Research Center and Biological
and Environmental Science and Engineering, King Abdullah University of Science
and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia

Email: takashi.gojobori@kaust.edu.sa

Present address at: ¹Department of Fisheries Technology, Bangladesh Agricultural
University, Mymensingh-2202, Bangladesh; ²Ofunato Fisheries
Promotion Center, Iwate Prefectural Government, Ofunato, Iwate 022-
8502, Japan.

Running title: Taxonomic profile in metagenomes of Ofunato Bay

ABSTRACT

The Ofunato Bay in Iwate Prefecture, Japan is a deep coastal bay located at the center of the Sanriku *Rias coast* and considered an economically and environmentally important asset. Here, we describe the first whole genome sequencing (WGS) study on the microbial community of the bay, where surface water samples were collected from three stations along its length to cover the entire bay; we preliminarily sequenced a 0.2 μm filter fraction among sequentially size-fractionated samples of 20.0, 5.0, 0.8 and 0.2 μm filters, targeting the free-living fraction only. From the 0.27–0.34 Gb WGS library, 0.9×10^6 – 1.2×10^6 reads from three sampling stations revealed 29 bacterial phyla (~80% of assigned reads), 3 archaeal phyla (~4%) and 59 eukaryotic phyla (~15%). Microbial diversity obtained from the WGS approach was compared with 16S rRNA gene results by mining WGS metagenomes, and we found similar estimates. The most frequently recovered bacterial sequences were Proteobacteria, predominantly comprised of 18.0–19.6% *Planktomarina* (Family Rhodobacteraceae) and 13.7–17.5% *Candidatus Pelagibacter* (Family Pelagibacterales). Other dominant bacterial genera, including *Polaribacter* (3.5–6.1%), *Flavobacterium* (1.8–2.6%), *Sphingobacterium* (1.4–1.6%) and *Cellulophaga* (1.4–2.0%), were members of Bacteroidetes and likely associated with the degradation and turnover of organic matter. The Marine Group I Archaea *Nitrosopumilus* was also detected. Remarkably, eukaryotic green alga *Bathycoccus*, *Ostreococcus* and *Micromonas* accounted for 8.8–15.2%, 3.6–4.9% and 2.1–3.1% of total read counts, respectively, highlighting their potential roles in the phytoplankton bloom after winter mixing.

Keywords

metagenomics

bacterial diversity

Ofunato Bay

whole genome sequencing

16S rRNA

ACCEPTED MANUSCRIPT

1. Introduction

Coastal bays are medium to large bodies of water, partially enclosed by land and connected to an ocean with a mouth. Their characteristic features vary depending on their mode of formation: from continental drift, river erosion to submergence of *Rias*. Among the ecosystems of all bay types, those formed by *Rias* are of importance due to their conspicuous terrace structure, primary productivity and biodiversity. The word *Ria* was coined by von Richthofen (1886) and is defined as mountainous-sided estuaries that eroded subaerially and former river valleys that have been drowned by sea level rise. Sanriku Coast in Japan is a typical *Rias Coast* and considered to be one of the most productive areas (Ishizaka, 1998). These waters are continuously under tidal influences by the cold *Oyashio* and the warm *Kuroshio* Currents, causing salinity, dissolved oxygen level and various nutrient components to vary with time, bringing in adequate inorganic nutrients. High productivity of the *Rias Coast* attracted aquaculture of many filter feeder species and brought about a substantial change to the extent of genetic diversity within communities. A combination of various agro-industrial activities, as well as human settlements and urbanization, has had a profound impact on the organisms that inhabit these waters and the ecosystem (Dugan et al., 2011).

The Ofunato Bay, an enclosed bay in Iwate Prefecture, is a deep coastal bay located at almost the center of the Sanriku *Rias Coast*. It possesses a complex physical structure and is famous for its complicated coastal line (Okaichi, 2004). The length and width are 7 and 2.5 km respectively, with an average depth of 10.3 m in the inner part and 38.5 m at the bay entrance. The bay entrance is protected by a tsunami breakwater that allows the passage of vessels 200 m in length. The chlorophyll *a* (chl-*a*) concentration has been reported to be high in the bay, especially during spring (JAXA, 2004), and attracts *Buoy-and-Rope type* scallop and oyster culture facilities (Komatsu et al., 2012). Annual production of the Japanese oyster (*Crassostrea gigas*) is approximately 300 mt wet weight (Miyazawa and Hayakawa, 1994).

Considering its economic and environmental importance, the Ofunato Bay has received much attention, and there has been continuous monitoring of its seawater quality parameters by organizations including the Japan Aerospace Exploration Agency and Iwate Fisheries Technology Center. The Ofunato Bay receives water from multiple sources, including the Sakari River, surface run-off from agriculture and industrial activities and the Sakari Wastewater Treatment Plant. Due to the calm nature of the bay, human settlements flourished near the bay in the past, and industrial activities, e.g., cement factories, were established. After 2011, new safety measures against tsunamis in the bay entrance and surrounding areas were taken by the government, and these changes are likely to increase human activities around the bay, thus leading to demographic changes (Cyranoski, 2012), which in turn may enhance run-off from households and wastewater plants and increase the internal load of nutrients in bay sediments. There have already been reports of frequent toxic dinoflagellate blooms in the Ofunato Bay (Ogata et al., 1982; Sekiguchi et al., 1989; Natsuike et al., 2014), primarily due to stratification and eutrophication (Ogata et al., 1982; Sellner et al., 2003). Water quality monitoring data over the last 20 years suggests dynamic changes have occurred in this valuable water resource, with depletion of dissolved oxygen near the bottom in summer (attributed to heavy biodeposition of organic matters from upper layers) and little exchange of seawater in strongly stratified bottom layers (Hayakawa et al., 2001). Several studies have provided insights on the planktonic populations of the bay (Ogata et al., 1982; Sakamoto et al., 1992), but detailed analyses on the associated microbial communities of the Ofunato Bay and their functions remain unreported. Additionally, the effect of human activities on the aquatic ecosystems (Oh et al., 2011; Ghai et al., 2012) requires close monitoring of the bay, particularly on its community structure, so that the sustainability of its fisheries and aquatic resources can be ensured.

Metagenomics has become the tool of choice that elucidates in-depth taxonomic profiles and functions of planktonic, bacterial, and viral communities while shedding light on interactions among community members (Kunin et al., 2008). Several metagenomic studies have demonstrated insights on the taxonomic diversity of microbial communities in marine and freshwater ecosystems from different parts of world (Venter et al., 2004; Rusch et al., 2007; Ghai et al., 2012; Grzymski et al., 2012; Ganesh et al., 2014; Brown et al., 2015) using direct metagenomic sequencing/whole genome sequencing (WGS) and amplicon-based metagenomic sequencing, with the former being able to retrieve microbial community composition more accurately (Shah et al., 2011; Logares et al., 2014). Additionally, the mining of WGS datasets for 16S rRNA is possible, which is free from common biases resulting from classical 16S rRNA amplification and cloning (Acinas et al., 2005; Temperton et al., 2009) used in taxonomic analyses, and this approach has been used successfully to elucidate microbial functions related to biogeochemical cycles of the ecosystem (Pfister et al., 2010). Due to the unavailability of information on microbial diversity and related functions in the Ofunato Bay, this study provides a comprehensive understanding of the ecosystem and function of microbes within the environment using a WGS approach and reports metagenomic analyses of the Sanriku *Rias Coast* for the first time.

2. Materials and methods

2.1 Sampling sites

The study area is in the Ofunato Bay, Iwate Prefecture, Japan. To comprehensively cover the total length of the bay, we selected three sampling locations for seawater samples: KSt.1 (141.73457E, 39.063370N; avg. depth 10.3 m), KSt. 2 (141.73245E, 39.044612N; avg. depth 25.3 m) and KSt. 3 (141.72820E, 39.019030N; avg. depth 38.5 m) (Fig. 1), where KSt. 1 and

3 are in the northeast (viz. inner part) and southwest (viz. bay entrance) part of the bay, while KSt. 2 is at the center of the bay and close to oyster culture facilities.

2.2 Sample collection and processing

Seawater sampling was carried out in January 2015 during the winter season, and samples were collected using a vertical water sampler at a 1 m depth from the surface. No specific permits were required for sampling, and the study did not involve any endangered or protected species. Approximately 8 L of seawater was pre-filtered through a 100- μ m filter to remove debris and was transported to the laboratory at Sanriku Coastal Education and Research Center, Kitasato University School of Marine Biosciences, Ofunato, Iwate Prefecture, under refrigerated condition ($\sim 4^{\circ}\text{C}$) for processing. Corresponding seawater quality parameters, including temperature, salinity and dissolved oxygen (DO), were also measured on site using an AAQ-RINKO AAQ176 (JFE Advantec Co. Ltd., Hyogo, Japan), and these properties are described in Supplementary Table S1. Of the collected seawater, 250 ml was used for nutrient and chlorophyll *a* (chl-*a*) measurements, and the remainder was filtered sequentially through 20- μ m pore sized Nylon Net Filters NY20 with a 47-mm diameter (Merck Millipore Ltd., Tullagreen, Ireland) and through 5-, 0.8- and 0.2- μ m pore sized Isopore Membrane Filters, 142 mm in diameter (Merck Millipore Ltd.), using a peristaltic pump (Masterflex L/S, Cole-Parmer International, IL, USA). To minimize disruption of soft-bodied protists and other picoplankton, we maintained a low pump speed of 30 – 40 rpm. After filtration, all filters were stored at -80°C until DNA extraction.

2.3 DNA extraction and shotgun metagenomic sequencing

Genomic DNA was extracted from 0.2- μ m filters using the PowerWater® DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) following the manufacturer's

recommendations. DNA concentrations were measured with a Qubit dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA, USA), read with a Qubit Fluorometer 2.0 (Invitrogen) and stored at -20°C until use. Extracted DNA aliquots were used for metagenomic shotgun sequencing. Five microliters of gDNA at a concentration of $0.2\text{ ng}/\mu\text{l}$ were taken, and DNA samples for metagenomics were prepared employing the NexteraXT DNA Sample Preparation Kit (Illumina Inc., San Diego, CA) and sequenced as paired-end reads using an Illumina MiSeq platform and a MiSeq Reagent Kit v3 (600 cycles) (Illumina Inc.).

2.4 Sequence processing and analyses

Illumina paired reads for each library were first joined by overlapping forward and reverse reads of the same DNA fragment (paired-end sequences) using the software FLASH (Magoc and Salzberg, 2011) with default parameters (overlap minimum, 10 nt; maximum allowed ratio between the number of mismatched basepairs and the overlap length, 0.25). Quality filtering of these WGS reads was then performed by removing reads $<50\text{ bp}$ and quality trimmed to Phred 20 using the Genomics Workbench (CLCbio, Cambridge, MA). BLASTn was performed locally on a server at the Sagamihara campus in the Kitasato University School of Marine Biosciences, Kanagawa Prefecture, using the quality- and size-filtered sequences against the NCBI-nt reference database. Taxonomic analysis to the species level was then performed using MEGAN v5.10.3 (Huson et al., 2007) after parsing the BLAST output. Comparative analysis in MEGAN was also performed after normalizing counts per the recommendations of the authors. The least common ancestor (LCA) assignment algorithm had the parameters: min support = 1; min score = 50; top percent = 10 and maximum expected $1.0\text{E} - 10$. In addition, a taxonomic analysis of WGS datasets targeting the rRNA gene sequences was conducted in the Meta Genome Rapid Annotation Subsystem Technology (MG-RAST) server. Data were first uploaded to MG-RAST and processed after

joining overlapping paired-ends reads and filtered using QC pipelines (Meyer et al. 2008) to remove replicated reads. Taxonomic analysis was then conducted by comparison against the Ribosomal Database Project (RDP) (Cole et al., 2009). Only alignments longer than 100 bp were considered for determining the minimum percent identity >95%. The best-named hit to the metagenomic rRNA sequence was used to assign the sequence to a high-level taxonomic group as described by Ghai et al. (2011).

3. Results and discussion

3.1 Characteristics of the metagenomic datasets

Using a shotgun metagenomics approach, we obtained sequence data from the 0.2- μ m filtered materials for all three sampling stations of the Ofunato Bay. Because our primary target was to conduct a taxonomic survey of free-living picoplankton population of the bay, we disregarded the particle-associated fractions of 5.0- and 0.8- μ m filter datasets and only included datasets for the 0.2- μ m filter fraction for this study. The sequence data contained 0.93 – 1.22 M sequences per sample (934,064 – 1,225,081 bp per sample), and the post-QC sequence length after joining overlapping forward and reverse reads ranged from 50 – 592 bp (mean length, 313 bp) (Table 1). These WGS datasets have been registered into the DDBJ database under the accession number DRA005744. The number of sequence reads resulting from WGS dataset mining for the 16S rRNA genes was 2,028 – 2,705, which accounted for 0.16– 0.17% of the corresponding WGS datasets (Table 2).

As the GC content of metagenomic dataset is a useful parameter to discriminate populations (Kunin et al. 2008) and is regarded as characteristic to a specific habitat, we calculated the GC signatures of WGS metagenomes of the Ofunato Bay for surface water collected from three stations. It was found that they exhibited a unimodal GC distribution with a single distinct peak at ~35%, typical for marine metagenomes (Ghai et al., 2011, 2012) (Fig. 2).

This pattern among three sampling stations also coincides with seawater salinity across the Ofunato Bay, which ranged from 32.9–33.6 psu during the sampling period, and preliminarily indicates closely related communities across three sampling stations (Ghai et al., 2012).

3.2 Taxonomic diversity of the Ofunato Bay

The metagenome sequences were largely dominated by bacteria, as shown by the analysis of both the shotgun metagenomics (69.5 – 79.7%) and the 16S rRNA (90.3 – 91.9%) reads recovered from the metagenomic data (Supplementary Figure S1). Archaeal sequences, on the other hand, accounted for 3.8 – 4.5% and 1.1 – 2.5% of the WGS and 16S rRNA reads, respectively. Interestingly, 1.1 – 1.4% WGS sequences were assigned to viruses, and it was also possible to identify eukaryotic sequences between 15–25% using the same approach. Few eukaryotes and no viruses were detected in the 16S reads. A similar pattern of organism domain assignments in shotgun metagenomic datasets has been reported by others (Debroas et al., 2009; Ghai et al., 2012; Jordaan and Bezuidenhout, 2013).

The taxonomic profiles of major bacterial phyla based on shotgun metagenomics and 16S rRNA reads recovered from the metagenome data were similar (Fig. 3). The most frequently recovered bacterial sequences detected by both the WGS and 16S approaches belonged to the Proteobacteria (Alpha, Beta and Gamma), followed by Bacteroidetes, collectively accounting for ~85% of all bacterial reads by both approaches (Supplementary Tables S2 and S3). Other phyla simultaneously detected include Firmicutes, Actinobacteria, Cyanobacteria and Tenericutes, consisting of ~1% by read counts. Although the classification of 16S rRNA datasets recovered from WGS is considered to be advantageous over that of WGS data itself (primarily because it is free from PCR-related biases), the former approach sometimes detects a lower number of bacterial genera in the metagenome (Brown et al., 2015). We observed major differences between these two approaches for 5 bacterial phyla (Aquificae,

Deferribacteres, Planctomycetes, Thermodesulfobacteria and Thermotogae) and 1 archaeal phylum (Thaumarchaeota) that were detected in WGS but absent from the 16S library (Fig. 3). The abundance of these differentially detected phyla in WGS datasets accounted for ~7% of all bacterial and archaeal reads, with Thaumarchaeota alone contributing ~6% in terms of read number.

3.2.1 *Proteobacteria*:

Proteobacteria formed more than half (60.3 – 61.8% of WGS reads and 68.5 – 71.2% of 16S) of the community in the Ofunato Bay (Fig. 3).

Alphaproteobacteria:

Both approaches confirmed the presence of large numbers of Alphaproteobacteria (50% of all WGS reads and 40% of all 16S) in the Ofunato Bay, and we observed that the most abundant taxa within this group was Rhodobacteraceae. Sakami et al. (2016) also reported a high abundance of Rhodobacteraceae in the Sendai Bay, which is located within 200 km of Ofunato Bay. In the WGS analysis, the most numerous genus in the Rhodobacteraceae was *Planktomarina* (Supplementary Table S2), which is regarded as the largest member of the marine *Roseobacter* clade-affiliated (RCA) cluster distributed widely in the temperate and sub-polar waters (Buchan et al., 2005; West et al., 2008; Giebel et al., 2009). *Planktomarina* has the unique ability to oxidize carbon monoxide (CO) as a source for energy and aerobic anoxygenic photosynthesis (AAnP) achieved by harvesting light via bacteriochlorophyll *a* (BChl_a) without CO₂ fixation. This free-living bacterioplankton plays a major role in the ocean biogeochemical cycle by processing multitude of organic compounds, particularly during phytoplankton blooms (Brinkhoff et al., 2008; Voget et al., 2015). Yamada et al. (2012) reported that in the Ofunato Bay, phytoplankton blooms generally occur during spring.

Our sampling was conducted during winter and in the present study; we observed no apparent phytoplankton bloom in the Ofunato Bay, and the abundance of *Planktomarina* ranged from 18.0–19.6%, which is in line with abundances reported by other studies (Giebele et al., 2011; Voget et al., 2015). The second most dominant bacterial genus identified by WGS analysis in the Ofunato Bay was *Candidatus Pelagibacter*, an abundant member of the SAR11 clade in the class Alphaproteobacteria. These are extreme oligotrophs and are unculturable bacteria. Many research groups tried to isolate these bacteria, and Rappé et al. (2012) obtained *Candidatus Pelagibacter ubique* as a culturable strain on a defined, artificial medium; it was found that *Candidatus Pelagibacter ubique* has a smaller genome (1,308,759 bp) than other free-living organisms with proteorhodopsin genes related to light-mediated proton pumps. In this study, we found that *Candidatus Pelagibacter ubique* comprised a substantial percentage (13.7–17.5%) of reads from the Ofunato Bay metagenomes (data not shown). Another species of this genus, *Candidatus Pelagibacter* sp. IMCC9063, was found in smaller numbers (data not shown). All members of the SAR11 lineage are the most dominant bacterioplankton in marine waters (Ghai et al., 2011). We found slight variations in the SAR11 lineage abundance among our three sampling stations, where they showed a gradual increase from KSt. 1 (~14%) through KSt. 3 (~18%), indicating a profound influence of the Sakari River near KSt. 1 and a higher influence of the Pacific Ocean near the mouth of the Bay at KSt. 3. Other alphaproteobacterial members represented by read counts greater than 1% in WGS analysis included *Bartonella*, *Octadecabacter* and *Roseobacter*. *Bartonella* are facultative intracellular bacteria, opportunistic pathogens and sometimes bacteremic for many warm-blooded animals, including rodents and ruminants (Jerris et al. 1996; Breitschwerdt et al. 2000; Dehio et al. 2004), suggesting them to be of terrestrial origin rather than marine. It is conceivable that *Bartonella* drifted down the river from terrestrial sources. The mountainous forests in the Iwate Prefecture in Japan are highly populated with Japanese sika deer (*Cervus*

nippon), which are reported to harbor several *Bartonella* species with infection rates as high as 41.5% (Sato et al., 2012). *Octadecabacter* species, on the other hand, are psychrophilic marine microorganisms collected from polar areas. Recently, *O. temperatus* SB1 was isolated and characterized as a nonpolar, mesophilic isolate of the genus *Octadecabacter*; growth occurred between 4 and 25 °C and was optimal at 20 °C (Gosink et al., 1997; Billerbeck, 2015). These results suggest that *Octadecabacter* species prefer low temperature in marine environments such as the Ofunato Bay. *Roseobacter* species are distributed in various marine habitats, including coastal to open oceans and sea ice to sea floor (Shiba, 1991). They have lived as free-living, particle-associated or in commensal relationships with marine phytoplankton, invertebrates, and vertebrates. As we used an isolation protocol to select for free-living organisms, only free-living *Roseobacter* was obtained in this study.

Betaproteobacteria:

Members of the Betaproteobacteria were detected in low abundance in the Ofunato Bay and are mainly regarded as freshwater bacteria that are relatively scarce in the marine environment (Ghai et al., 2012). The most abundant member of this group was *Candidatus Kinetoplastibacterium* followed by *Polynucleobacter* (Supplementary Table S2). *Candidatus Kinetoplastibacterium* are bacterial endosymbionts in insect-infecting trypanosomatid flagellates (Alves et al., 2013). With their smaller size and reduced genome, they adapted to their hosts and supplement the host metabolism with essential compounds including heme, nucleotides and amino acids (de Souza and Motta, 1999; Alves et al., 2011). Although these bacteria have been reported to be unculturable to date (de Souza and Motta, 1999), many Betaproteobacteria have culturable representatives (Hahn et al., 2010; Segawa et al., 2011), and the availability of a small number of free-living *Candidatus Kinetoplastibacterium* in the water fraction of a semi-closed bay such as the Ofunato Bay points to the possibility that

there are viable communities on the surface waters. *Polynucleobacter*, on the other hand, are free-living bacteria and are abundant in freshwater systems such as lakes, ponds, and streams (Hahn et al., 2009). It appears that these microorganisms have been washed away in the Sakari River along with the rainwater, where like the free-living Alphaproteobacteria, they showed a relatively higher abundance at KSt. 1 and lower abundance at KSt. 3.

Deltaproteobacteria:

Members of the Deltaproteobacteria were also detected in low abundance in the Ofunato Bay, and the genus *Bacteriovorax* is present in only 0.1% of the WGS reads (Supplementary Table S2). The genus *Bacteriovorax* is a predator that feeds on larger bacteria (Baer et al., 2000). *Bacteriovorax* has similar morphological and lifestyle characteristics as *Bdellovibrio* (Baer et al., 2004), suggesting that the genus *Bacteriovorax* is one of the consumers in the food chain of the Ofunato Bay.

Epsilonproteobacteria:

Members of Epsilonproteobacteria were also detected in low abundance in the Ofunato Bay, with the genus *Arcobacter* representing approximately 0.3% of all WGS reads. The genus *Arcobacter* inhabits various environments, and some species can be human and animal pathogens (Miller et al., 2007). They are aerotolerant like *Campylobacter*, but *Arcobacter* species can grow at lower temperatures than *Campylobacter*. It seems that *Arcobacter* is adapted for life in the Ofunato Bay.

Gammaproteobacteria:

The most dominant Gammaproteobacteria member detected in the Ofunato Bay was *Buchnera*, represented by read counts greater than 0.5% in WGS analysis (Supplementary

Table S2). *Buchnera* is a close relative of *Escherichia coli* (Baumann et al., 2000) and a primary endosymbiont of most aphids (Munson et al. 1991; Nakabachi et al., 2005), where aphids developed bacteriocyte cells to host *Buchnera*. Reproduction of *Buchnera* is host dependent, being transmitted maternally to eggs and embryos through host generations, and is contained within the bacteriomes (Shigenobu et al., 2000). Considering the absolute dependence of *Buchnera* on their host, and our findings of their slightly higher relative abundance in KSt. 1 compared to other two sampling stations, we conclude that they originated from the surrounding mountainous areas of the Ofunato Bay, drifted down the river and found their way into the bay. Other members of this group, represented by read counts below 0.5% relative abundance, included *Shewanella*, *Alteromonas* and *Francisella*. *Shewanella* are distributed in the tropical marine environment and can utilize electron acceptors under anaerobic conditions (Kim et al., 2007). *Alteromonas* are also typical marine bacteria found in the coastal waters (Baumann et al., 1972) and considered as decomposers in the food chain. These group members are generally available in nutrient-rich aquatic environments (Hicks et al., 2000; Nealson and Scott, 2003), and their availability in the Ofunato Bay appears to be linked with high nutrient content and primary productivity. *Francisella*, on the other hand, are iron-dependent facultative intracellular parasites of macrophages (Allen et al., 2003; Sullivan et al., 2006). Their presence in the Ofunato Bay suggests that they were leaked out from macrophages, as we used bacteria isolation protocol to select for free-living organisms only.

3.2.2 Bacteroidetes:

Bacteroidetes were the second most abundant bacterial phylum detected in the Ofunato Bay, and the taxonomy assignment included a diverse clade including *Polaribacter* (3.5–6.1%), *Flavobacterium* (1.8–2.6%), *Sphingobacterium* (1.4–1.6%) and *Cellulophaga* (1.4–2.0%).

Polaribacter are non-motile rods isolated from sea ice, water and soil from temperate regions (Fukui et al., 2013; Xing et al., 2015) and belong to the family Flavobacteriaceae. These bacteria have been reported to be prevalent in the Arctic Sea, and their abundance in the Ofunato Bay suggests that *Polaribacter* drifted here by the *Oyashio* current. Indeed, their abundance was higher near the mouth of the bay (6.1% of assigned WGS reads at KSt. 3) and probably linked to enhanced mixing of seawater between the inside and outside of the bay (Yamada et al., 2017), which gradually declined towards the inner part (4.7% at KSt. 2 and 3.5% at KSt. 1). BLASTn and MEGAN analyses revealed *Polaribacter* sp. MED152 to be the major species of this genus that possesses proteorhodopsin and genes for attachment and polymer degradation, suggesting their ability to utilize solar energy and maintain survival even under low-nutrient conditions (Gosink et al., 1998; González et al., 2008). Other genera of the phylum Bacteroidetes detected in the Ofunato Bay perform variety of important ecological functions, including degradation of proteins and polysaccharides and turnover of organic matters (Cottrel and Kirchman, 2000).

3.2.3 Actinobacteria:

Actinobacteria are generally highly abundant in both marine and freshwater communities. In the present study, we observed that 1.5 – 2.1% of WGS reads and 0.8 – 0.9% of 16S reads were assigned to Actinobacteria. The most dominant genus in this group was *Candidatus Actinomarina*, the smallest known free-living cells described to date (Ghai et al., 2013). They are widely distributed in the photic zone of oceans in temperate and tropical regions, and with the presence of rhodopsins, perform photoheterotrophic metabolism to thrive in the complex aquatic environment (Ghai et al., 2013).

3.2.4 Cyanobacteria:

Cyanobacteria generally forms a significant part of the marine picoplankton community, particularly on the surface, and in our study, 0.8–1.2% of WGS reads were assigned to the Cyanobacteria. Taxonomy based on 16S reads, however, revealed little or no Cyanobacteria in Ofunato Bay. This is possibly related to the fact that half of the bacterial 16S sequences (47.5–56.3%) remained unclassified using this approach. The WGS metagenomics data strongly indicated the presence of Cyanobacteria, where the top organism identified was *Synechococcus*, followed by *Prochlorococcus* and *Nostoc* (Supplementary Table S3). In aquatic environments, *Synechococcus* and *Prochlorococcus* are important contributors of primary production (Palenik et al., 2003; Ghai et al., 2012; Brown et al., 2015), and in environments where *Synechococcus* and *Prochlorococcus* are simultaneously present, the former alone contributes to the primary productivity (Ghai et al., 2012). These results suggest that members of the Cyanobacteria are important producers in the food cycle of Ofunato Bay.

3.2.5 Archaea:

From the metagenomic datasets, it was possible to identify archaeal sequences using both WGS and 16S analyses, and we found that they accounted for 3.8–4.5% and 1.1–2.5% of all assigned reads in WGS and 16S approaches, respectively (Supplementary Table S2). *Nitrosopumilus* was the most dominant genus in the group (4.2 – 4.8% in the WGS and 1.1 – 2.6% in the 16S library) (Supplementary Tables S2 and S3), which is regarded as an important member of the Marine Group I Archaea and is capable of oxidizing ammonia at even extremely low concentrations (Martens-Habben et al. 2009; Walker et al. 2010). Ammonia concentration varies in natural environments, and open oceans have relatively low amounts. Since ammonia-oxidizing Archaea such as *Nitrosopumilus* are able to thrive under limited nutritional conditions, their mechanisms of highly specific ammonia oxidation rates at

very low ammonia concentrations and autotrophy are thought to play an important role in global biogeochemical cycles (Konneke et al., 2005; Schleper et al., 2005; Francis et al., 2005). Indeed, the ammonia concentrations of the three sampling stations in the Ofunato Bay ranged from 1.2 – 1.5 $\mu\text{mol/l}$ at the time of sampling (Supplementary Table S1), which is equivalent to those of open oceans (Martens-Habben et al., 2009). As we observed that the relative abundance of *Nitrosopumilus* did not vary significantly across the three sampling stations under investigation and ammonia concentrations had a steady range below 10 $\mu\text{mol/l}$, we predict that ammonia concentrations determine their growth and abundance, as reported previously for marine and coastal sediments (Park et al., 2010; Verhamme et al., 2011). Other Archaea genera detected at a read cutoff of ≥ 20 included *Methanococcus* and *Methanobrevibacter* (Supplementary Table S2), both of which are methanogens that recycle hydrogen by combining it with carbon dioxide to produce methane (Haney et al., 1999; Samuel et al., 2007).

3.2.6 Eukaryotes:

Eukaryotes accounted for 8.8 – 15.2% of the WGS and 2.2 – 3.9% of the 16S metagenomic reads from the three sampling stations in the Ofunato Bay (Fig. 3). Taxonomic profiles for eukaryotes using both approaches were similar. It was observed that *Bathycoccus*, *Ostreococcus* and *Micromonas* were the major eukaryotic genera having read counts greater than 1% of total reads in WGS datasets, while *Dinophysis* was the major genera in the 16S datasets. In the WGS datasets, the three photosynthetic picoplankton genera formed the majority (~85%) of the eukaryotic community in Ofunato Bay and are considered to be the major producers in the oceanic food webs, contributing up to as much as 68% of the total primary production (Li et al., 1993; Vaulot et al., 2012) linked primarily to their larger biomass (cell size $< 2 - 3 \mu\text{m}$) and higher growth rates (Worden et al., 2004). *Bathycoccus*,

Ostreococcus and *Micromonas* have been reported to form blooms in temperate waters, particularly during and just after the deep winter mixing period (Treusch et al., 2012). In the Ofunato Bay, water mixing occurs during November through January (Yamada et al., 2017), and as our samples were collected at the end of mixing period, further analyses using samples collected after that period may confirm phytoplankton blooms and related changes in the planktonic community.

3.2.7 Viruses:

Viruses generally represent a small portion of the marine community, and in the present study, we observed *Prasinovirus* to be the most abundant viral genus, accounting for 0.6 – 0.8% of WGS reads. *Prasinovirus* are known as phycodnaviruses, as they use algae as hosts (Marin and Melkonian, 2010), and they have been shown to have important roles in microbial mortality and marine biogeochemical cycles (Clerissi et al., 2015). The abundance of green alga in the Ofunato Bay was substantially high at the time of sampling (Supplementary Table S3), and we observed a positive correlation between *Prasinovirus* abundance and green alga (data not shown). In the WGS library, a small number of unclassified Myoviridae, Podoviridae and Siphoviridae were detected that infect bacteria and Archaea, including highly abundant *Synechococcus* in the Ofunato Bay (Lavigne et al., 2009; Ghai et al., 2012; Kleppen et al., 2012; Bebeacua et al. 2013).

4. Conclusions

In the present study, we obtained genus-level taxonomic profiles of the free-living microbial community in the Ofunato Bay using a WGS approach. As expected, Proteobacteria (Alpha, Beta and Gamma) were the most abundant phyla in the Ofunato Bay, followed by Bacteroidetes, Firmicutes, Actinobacteria, Cyanobacteria and Tenericutes. We detected a wide range of heterotrophic bacteria, including *Planktomarina*, *Candidatus* Pelagibacter, *Polaribacter*, *Flavobacterium*, *Sphingobacterium*, *Candidatus* Actinomarina and Cellulophaga, Archaea including *Nitrosopumilus* and eukaryotic algae including *Bathycoccus*, *Ostreococcus* and *Micromonas*. Most of these bacteria are considered significant in oceans and marine ecology. *Bartonella*, *Candidatus* Kinetoplastibacterium, and *Buchnera* DNAs were found in seawater samples of the Ofunato Bay in this study; these results are surprising because the habitats of these microbes are typically terrestrial areas rather than marine areas. We also found various bacterial DNAs that are symbiotic, parasitic and/or pathogenic to humans, animals, plants and insects in terrestrial areas. It is conceivable that these microbes drifted down the river to the Ofunato Bay directly or indirectly. It appears that the high detection frequency of these microbes is due to the enclosed character of the Ofunato Bay. It is not clear if they are still alive in seawater, but the presence of their DNA in seawater contributes to the biodiversity of DNA resources in the Ofunato Bay.

Acknowledgments

We thank Dr. Tomoko Sakami, Tohoku National Fisheries Research Institute, Japan Fisheries Research and Education Agency, for helping establish our sampling and DNA extraction systems. The research reported in this publication was supported in part by the King Abdullah University of Science and Technology (KAUST), under award number URF/1/1976.

References

- Acinas, S.G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., Polz, M.F., 2005. PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Appl. Environ. Microbiol.* 71, 8966–8969.
- Allen, L. A. (2003). Mechanisms of pathogenesis: evasion of killing by polymorphonuclear leukocytes. *Microbes Infect.* 5, 1329–35.
- Alves, J.M., Voegtly L., Matveyev A.V., Lara A.M., da Silva F.M., Serrano M.G., Buck G.A., Teixeira M.M., Camargo E.P., 2011. Identification and phylogenetic analysis of hemesynthesis genes in trypanosomatids and their bacterial endosymbionts. *PLoS One* 6, e23518.
- Alves, J.M., Serrano M.G., Maia da Silva F., Voegtly L.J., Matveyev A.V., Teixeira M.M., Camargo E.P., Buck G.A., 2013. Genome evolution and phylogenomic analysis of *Candidatus Kinetoplastibacterium*, the Betaproteobacterial endosymbionts of *Strigomonas* and *Angomonas*. *Genome Biol. Evol.* 5, 338–350.
- Baer, M.L., Ravel, J., Pineiro, S.A., Guether-Borg, D., Williams, H.N., 2004. Reclassification of salt-water *Bdellovibrio* sp. as *Bacteriovorax marinus* sp. nov. and *Bacteriovorax litoralis* sp. nov. *Int. J. Syst. Evol. Microbiol.* 54, 1011–1016.
- Baer, M., Ravel J., Chun J., Hill R., Williams H., 2000. A proposal for the reclassification of *Bdellovibrio stolpii* and *Bdellovibriostarrii* into a new genus, *Bacteriovorax* gen. nov. as *Bacteriovorax stolpii* comb. nov. and *Bacteriovorax starrii* comb. nov., respectively. *Int. J. Syst. Evol. Microbiol.* 50, 219–224.
- Baumann, L., Baumann P., Mandel M., Allen R.D., 1972. Taxonomy of aerobic marine eubacteria. *J. Bacteriol.* 110, 402–429.
- Baumann, P., Moran, N. A., Baumann, L., 2000. *The Prokaryotes* (Dworkin, M., ed.) Springer, New York.

- Bebeacua, C., Lai, L., Vegge, C. S., Brøndsted, L., van Heel, M., Veessler, D., Cambillau, C.e,2013. Visualizing a complete Siphoviridae member by single-particle electron microscopy: the structure of Lactococcal phage TP901-1. *Virology* 87, 1061–1068.
- Bergey's Manual of Systematic Bacteriology, 2001. 2nd ed., vol. 1 (D.R. Boone, R.W. Castenholz, eds.), Springer-Verlag, New York, pp. 465–466.
- Billerbeck, S., Orchard, J.B., Tindall, B.J., Brinkhoff, T., Simon, M., 2015.Description of *Octadecabacter temperatus* sp. nov., isolated from the southern North Sea, emended description of the genus *Octadecabacter* and its species and reclassification of *Octadecabacter jejudonensis* Park & Yoon, 2014 as *Pseudooctadecabacter jejudonensis* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* 65, 1967–1974.
- Breitschwerdt E.B., Kordick D.L., 2000. *Bartonella* infection in animals: carriership, reservoir potential, pathogenicity, and zoonotic potential for human infection. *Clin. Microbiol. Rev.* 13, 428–438.
- Brinkhoff, T., Giebel, H.-A., Simon, M.,2008. Diversity, ecology,and genomics of the *Roseobacter* clade: a short overview. *Arch. Microbiol.* 189, 531–539.
- Clerissi, C., Desdevises, Y., Romac, S., Audic, S., de Vargas, C., Acinas, S. G., Casotti, R., Poulain, J., Wincker, P., Hingamp, P., Ogata, H., Grimsley, N., 2015. Deep sequencing of amplified Prasinovirus and host green algal genes from an Indian Ocean transect reveals interacting trophic dependencies and new genotypes. *Environ. Microbiol. Rep.* 7, 979–989.
- Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Marsh, T., Garrity, G.M., Tiedje, J.M., 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 37, D141–145.

- Cottrell, M. T., Kirchman, D. L., 2000. Natural assemblages of marine Proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular weight dissolved organic matter. *Appl. Environ. Microbiol.* 66, 1692–1697.
- Cyranoski, D., 2012. Rebuilding Japan: after the deluge. *Nature* 483, 141–143.
- de Souza W., Motta M. C., 1999. Endosymbiosis in protozoa of the Trypanosomatidae family. *FEMS Microbiol. Lett.* 173, 1–8.
- Dehio, C., Sauder, U., Hiestand, R., 2004. Isolation of *Bartonella schoenbuchensis* from *Lipoptena cervi*, a blood-sucking arthropod causing deer ked dermatitis. *J. Clin. Microbiol.* 42,5320–5323.
- Dugan, J.E., Airoidi, L., Chapman, M.G., Walker, S.J., Schlacher, T., Wolanski, E., McLusky, D., 2011. Estuarine and coastal structures: environmental effects, a focus on shore and nearshore structures. In: *Treatise on Estuarine and Coastal Science* (Wolanski E., McLusky D. ,eds). Academic Press, Waltham,pp.17–41.
- Dupont, C. L., Rusch, D. B., Yooseph, S., Lombardo, M.-J., Richter,R.A., Valas, R., Novotny,M., Yee-Greenbaum,J., Selengut,J.D.,Haft, D.H.,Halpern, A.L., Lasken, R.S., Nealson, K., Friedman,R., Venter, J.C., 2012. Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J.* 6, 1186–1199.
- Francis, C. A., Roberts, K. J., Beman, J. M., Santoro, A. E., Oakley, B. B., 2005. Ubiquity and diversity of ammonia-oxidizing Archaea in water columns and sediments of the ocean. *Proc. Natl. Acad. Sci. USA* 102, 14683–14688.
- Fukatsu, T., Ishikawa, H. 1992. A novel eukaryotic extracellular symbiont in an aphid, *Asteopteryx styraci* (Homoptera, Aphididae, Hormaphidinae), *J. Insect Physiol.* 38, 765–773.
- Fukui Y., Abe M., Kobayashi M., Saito H., Oikawa H., Yano Y., Satomi M.,2013. *Polaribacter porphyrae* sp. nov., isolated from the red alga *Porphyra yezoensis*, and

- emended descriptions of the genus *Polaribacter* and two *Polaribacter* species. Int. J. Syst. Evol. Microbiol. 63, 1665–1672.
- Ghai, R., Mizuno, C. M., Picazo, A., Camacho, A., Rodriguez-Valera, F., 2013. Metagenomics uncovers a new group of low GC and ultra-small marine Actinobacteria. Sci. Rep. 3, 2471.
- Giebel, H-A., Kalhoefer, D., Lemke, A., Thole, S., Gahl-Janssen, R., Simon, M., Brinkhoff, T., 2011. Distribution of *Roseobacter* RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. ISME J. 5, 8–19.
- Giebel, H. A., Brinkhoff, T., Zwisler, W., Selje, N., Simon, M., 2009. Distribution of *Roseobacter* RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. Environ. Microbiol. 11, 2164–2178.
- González, J. M., Fernández-Gómez, B., Fernández-Guerra, A., Gómez-Consarnau, L., Sánchez, O., Coll-Lladó, M., Del Campo, J., Escudero, L., Rodríguez-Martínez, R., Alonso-Sáez, L., Latasa, M., Paulsen, I., Nedashkovskaya, O., Lekunberri, I., Pinhassi, J., Pedrós-Alió, C., 2008. Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). Proc. Natl. Acad. Sci. USA 105, 8724–8729.
- Gosink, J.J., Herwig, R.P., Staley, J.T. 1997. *Octadecabacterarcticus* gen. nov., sp. nov., and *O. antarcticus*, sp. nov., nonpigmented, psychrophilic gas vacuolate bacteria from polar sea ice and water. Syst. Appl. Microbiol. 20, 356–365.
- Gosink, J.J., Woese, C.R., Staley, J.T., 1998. *Polaribacter* gen. nov., with three new species, *P. irgensii* sp. nov., *P. franzmannii* sp. nov. and *P. filamentus* sp. nov., gas vacuolate polar marine bacteria of the Cytophaga-Flavobacterium-Bacteroides group and reclassification of *Flectobacillus glomeratus* as *Polaribacter glomeratus* comb. nov. Int. J. Syst. Bacteriol. 48, 223–35.

- Grzymiski, J. J., Riesenfeld, C. S., Williams, T. J., Dussaq, A. M., Ducklow, H. W., Erickson, M., Cavicchioli, R., Murray, A. E., 2012. A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *ISME J.* 6, 1901–1915.
- Hahn, M. W., Kasalicky, V., Jezbera, J., Brandt, U., Simek, K., 2010. *Limnohabitans australis* sp. nov., isolated from a freshwater pond, and emended description of the genus *Limnohabitans*. *Int. J. Syst. Evol. Microbiol.* 60, 2946–2950.
- Hahn, M.W., E. Lang, U. Brandt, Q. L. Wu, Scheuerl, T., 2009. Emended description of the genus *Polynucleobacter* and the species *Polynucleobacter necessarius* and proposal of two subspecies, *P. necessarius* subsp. *Necessarius* subsp. nov. and *P. necessarius* subsp. *Asymbioticus* subsp. nov. *Int. J. Syst. Evol. Microbiol.* 59, 2002–2009.
- Haney, P.J., Badger, J.H., Buldak, G.L., Reich, C.I., Woese, C.R., Olsen, G.J., 1999. Thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species. *Proc. Natl. Acad. Sci. USA* 96, 3578–3583.
- Hayakawa, Y., Kobayashi, M., Hayashizaki, K., 2002. Environmental monitoring and data utilization for estimation of sedimentation flux from oyster mariculture in Ofunato estuary, Japan. *Fish. Sci.* 68 (suppl. I), 550–553.
- Hayakawa, Y., Kobayashi, M., Izawa, M., 2001. Sedimentation flux from mariculture of oyster (*Crassostrea gigas*) in Ofunato estuary, Japan. *ICES J. Mar. Sci.* 58, 435–444.
- Hicks, C.L., Kinoshita, R., Ladds, P.W., 2000. Pathology of melioidosis in captive marine mammals. *Aust. Vet. J.* 78, 193–195.
- Jerris, R.C, Regnery, R.L., 1996. Will the real agent of cat-scratch disease please stand up? *Annu. Rev. Microbiol.* 50, 707–25.

- Johansen, J.E., Nielsen, P., Sjöholm, C., 1999. Description of *Cellulophaga balticagen.* nov., sp. nov. and *Cellulophaga fucicola* gen. nov., sp. nov. and reclassification of [*Cytophaga*] *lytica* to *Cellulophaga lytica* gen. nov., comb. nov. Int. J. Syst. Bacteriol. 49, 1231–1240.
- Jones, R. D., Amador, J. A., 1993. Methane and carbon monoxide production, oxidation, and turnover times in the Caribbean Sea as influenced by the Orinoco River. J. Geophys. Res. 98, 2353–2359.
- Kim, D., Baik, K., Kim, M., Jung, B., Shin, T., Chung, G., Rhee, M., Seong, C., 2007. *Shewanellahaliotis* sp. nov., isolated from the gut microflora of abalone, *Haliotis discus hannai*, Int. J. Syst. Evol. Microbiol. 57, 2926–2931.
- Kleppen, H. P., Holo, H., Jeon, S.R., Nes, I. F., Yoon, S.S., 2012. Novel Podoviridae family Bacteriophage infecting *Weissella cibaria* isolated from kimchi. Appl. Environ. Microbiol. 78, 7299–7308.
- Konneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B., Stahl, D.A., 2005. Isolation of an autotrophic ammonia-oxidizing marine Archaeon. Nature 437, 543–546.
- Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K., Hugenholtz, P., 2008. A bioinformatician's guide to metagenomics. Microbiol. Mol. Biol. Rev. 72, 557–578.
- Lavigne, R., Darius, P., Summer, E. J., Seto, D., Mahadevan, P., Nilsson, A. S., Ackermann, H. W., Kropinski, A. M., 2009. Classification of Myoviridae bacteriophages using protein sequence similarity. BMC Microbiol. 9, 224.
- Li, W.K.W., Dickie, P.M., Harrison, W.G., Irwin, B.D., 1993. Biomass and production of bacteria and phytoplankton during the spring bloom in the western North-Atlantic Ocean. Deep-Sea Res. II-Topical Stud. Oceanograph. 40, 307–327.

- Logares, R., Sunagawa, S., Salazar, G., Cornejo-Castillo, F. M., Ferrera, I., Sarmiento, H., Hingamp, P., Ogata, H., de Vargas, C., Lima-Mendez, G., Raes, J., Poulain, J., Jaillon, O., Wincker, P., Kandels-Lewis, S., Karsenti, E., Bork, P., Acinas, S. G., 2014. Metagenomic 16SrDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environ. Microbiol.* 16, 2659–2671.
- Lozupone, C.A., Knight, R. 2007. Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci USA* 104, 11436–11440.
- Magoc, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963.
- Marie, D., Zhu, F., Balague, V., Ras, J., Vaulot, D. 2006. Eukaryotic picoplankton communities of the Mediterranean Sea in summer assessed by molecular approaches (DGGE, TTGE, QPCR). *FEMS Microbiol. Ecol.* 55, 403–415.
- Martens-Habbena, W., Berube, P.M., Urakawa, H., de la Torre, J.R., Stahl, D.A., 2009. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and bacteria. *Nature* 461, 976–979.
- McCarren, J., Becker, J.W., Repeta, D.J., Shi, Y., Young, C.R., Malmstrom, R.R., Chisholm, S.W., DeLong, E.F., 2010. Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc. Natl. Acad. Sci. USA* 107, 16420–16427.
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., Edwards, R.A., 2008. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9, 386–393.

- Miller, W. G., Parker, C. T., Rubenfield, M., Mendz, G. L., Wösten, M. M. S. M., Ussery, D. W., Stolz, J. F., Binnewies, T. T., Hallin, P. F., Wang, G., Malek, J. A., Rogosin, A., Stanker, L. H., Mandrell, R. E., 2007. The complete genome sequence and analysis of the *Epsilonproteobacterium Arcobacter butzleri*. PLoS ONE2, e1358.
- Miyazawa, K., Hayakawa, Y., 1994. A large-scale structure and the environmental capacity for aquaculture – effects of an artificial structure (breakwater at the entrance of the Ofunato estuary) on aquaculture and water qualities in the estuary. Bull. Coast. Oceanogr. 32, 29–38 (in Japanese with English abstract).
- Munn, C., 2011. Marine Microbiology: Ecology and Applications, 2nd ed., Garland Science.
- Munson, M. A., Clark, M. A., Baumann, L., Moran, N. A., Voegtlin, D. J., Campbell, B., 1991. Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. J. Bacteriol. 173, 6321–6324.
- Nakabachi, A., Shigenobu, S., Sakazume, N., Shiraki, T., Hayashizaki, Y., Carninci, P., Ishikawa, H., Kudo, T., Fukatsu, T., 2005. Transcriptome analysis of the aphid bacteriocyte, the symbiotic host cell that harbors an endocellular mutualistic bacterium, *Buchnera*. Proc. Natl. Acad. Sci. USA 102, 5477–5482.
- Natsukea, M., Kanamoria, M., Babad, K., Moribea, K., Yamaguchia, A., Imai I., 2014. Changes in abundances of *Alexandrium tamarense* resting cysts after the tsunami caused by the Great East Japan Earthquake in Funka Bay, Hokkaido, Japan. Harmful Algae 39, 271–279.
- Nealson, K.H., Scott, J., 2003. The prokaryotes: an evolving electronic resource for the microbiological community (Dworkin, E.A., ed.). Springer-Verlag, New York.
- Ogata, T., Kodama, M., Fukuyo, Y., Inoue, T., Kamiya, H., Matsuura, F., Sekiguchi, K., Watanabe, S., 1982. The occurrence of *Protogonyaulax* spp. in Ofunato Bay, in

- association with the toxification of the scallop *Patinopecten yessoensis*. Nippon Suisan Gakkaishi 48, 563-566.
- Oh, S., Caro-Quintero, A., Tsementzi, D., DeLeon-Rodriguez, N., Luo, C., Poretsky, R., Konstantinidis, K.T., 2011. Metagenomic insights into the evolution, function, and complexity of the planktonic microbial community of Lake Lanier, a temperate freshwater ecosystem. Applied. Environ. Microbiol. 77, 6000–6011.
- Okaichi, T., 2004. Red-tide species and the environmental conditions. In: Red Tides (Okaichi, T., ed.). Terra Scientific Publishing Company, Tokyo, Japan, and Kluwer Academic Publishers, London, Boston, p. 367.
- Palenik, B., Brahmsha, B., Larimer, F. W., Land, M., Hauser, L., Chain, P., Lamerdin, J., Regala, W., Allen, E. E., McCarren, J., Paulsen, I., Dufresne, A., Partensky, F., Webb, E. A., Waterbury, J., 2003. The genome of a motile marine *Synechococcus*. Nature 424, 1037–1042.
- Park, B.J., Park, S.J., Yoon, D.N., Schouten, S., Sinninghe Damsté, J. S., Rhee, S.K., 2010. Cultivation of autotrophic ammonia-oxidizing Archaea from marine sediments in coculture with sulfur-oxidizing bacteria. Appl. Environ. Microbiol. 76, 7575–7587.
- Potts, M., 1997. Etymology of the genus name Nostoc (Cyanobacteria). Int. J. Syst. Bacteriol. 47, 584.
- Rappé, M.S., Connon, S.A., Vergin, K.L., Giovannoni, S.J., 2002. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature 418, 630–633.
- Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., 2007. The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through eastern tropical Pacific. PLoS Biol. 5, e77.

- Sakami, T., Watanabe, T., Kakehi, S., Taniuchi, Y., Kuwata, A., 2016. Spatial variation of bacterial community composition at the expiry of spring phytoplankton bloom in Sendai Bay, Japan. *Gene* 576, 610–617.
- Sakamoto, S., Ogata, T., Sato S., Kodama, M., Takeuchi, T., 1992. Causative organism of paralytic shellfish toxins other than toxic dinoflagellates. *Mar. Ecol. Progress Ser.* 89, 229–235.
- Samuel, B.S., Hansen, E.E., Manchester, J.K., Coutinho, P.M., Henrissat, B., Fulton, R., Latreille, P., Kim, K., Wilson, R.K., Gordon, J.I., 2007. Genomic and metabolic adaptations of *Methanobrevibacter smithii* to the human gut. *Proc. Natl. Acad. Sci. USA* 104, 10643–10648
- Sato, S., Kabeya, H., Yamazaki, M., Takeno, S., Suzuki, K., Kobayashi, S., Souma, K., Masuko, T., Chomel, B.B., Maruyama, S., 2012. Prevalence and genetic diversity of *Bartonella* species in sika deer (*Cervus nippon*) in Japan. *Comp. Immunol. Microbiol. Infect Dis.* 35, 575–581.
- Schleper, C., Jurgens, G., Jonscheit, M., 2005. Genomic studies of uncultivated Archaea. *Nature Rev. Microbiol.* 3, 479–488.
- Sekiguchi, K., Inoguchi, N., Shimizu, M., Saito, S., Watanabe, S., Ogata, T., Kodama, M., Fukuyo, Y., 1989. Occurrence of *Protogonyaulax tamarensis* and shellfish toxicity in Ofunato Bay from 1980–1986. In: *Red Tides: Biology, Environmental Science, and Toxicology* (Okaichi, T., Anderson, D. M., Nernoto, T., eds.). Elsevier, New York, pp. 399–402.
- Sellner, K. G., Doucette, G. J., Kirkpatrick, G. J., 2003. Harmful algal blooms: causes, impacts and detection. *J. Ind. Microbiol. Biotechnol.* 30, 383–406.

- Shah, N., Tang, H., Doak, T. G., Ye, Y., 2011. Comparing bacterial communities inferred from 16SrRNA gene sequencing and shotgun metagenomics. *Pac. Symp. Biocomput.* 16, 165–176.
- Shi, Y., McCarren, J., DeLong, E.F., 2012. Transcriptional responses of surface water marine microbial assemblages to deep-sea water amendment. *Environ. Microbiol.* 14,191–206.
- Shiba, T., 1991. *Roseobacter litoralis* gen. nov., sp. nov., and *Roseobacter denitrificans* sp. nov., aerobic pink-pigmented bacteria which contain bacteriochlorophyll *a*. *Syst. Appl. Microbiol.* 14, 140-145.
- Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., Ishikawa H., 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* 407, 81-86.
- Sullivan, J. T., Jeffery, E. F., Shannon, J. D., Ramakrishnan, G., 2006. Characterization of the siderophore of *Francisella tularensis* and role of fslA in siderophore production. *J. Bacteriol.* 188, 3785–3795.
- Temperton, B., Field, D., Oliver, A., Tiwari, B., Muhling, M., Joint, I., Gilbert, J. A., 2009. Bias in assessments of marine microbial biodiversity in fosmid libraries as evaluated by pyrosequencing. *ISME J.* 3, 792–796.
- Treusch, A.H., Demir-Hilton, E., Vergin, K.L., Worden, A.Z., Carlson, C.A., Donatz, M.G., Burton, R.M., Giovannoni, S.J., 2012. Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. *ISME J.* 6,4 81–492.
- Treusch, A.H., Vergin, K.L., Finlay, L.A., Donatz, M.G., Burton, R.M., Carlson, C.A., Giovannoni, S.J., 2009. Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME J.* 3, 1148–1163.

- Vaulot, D., Lepère, C., Toulza, E., De la Iglesia, R., Poulain, J., Gaboyer, F., Moreau, H., Vandepoele, K., Ulloa, O., Gavory, F., Piganeau, G., 2012. Metagenomes of the picoalga *Bathycoccus* from the Chile coastal upwelling. PLoS ONE 7, e39648.
- Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.H., Smith, H.O., 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science 304, 66–74.
- Verhamme, D.T., Prosser, J.I., Nicol, G.W., 2011. Ammonia concentration determines differential growth of ammonia-oxidising Archaea and bacteria in soil microcosms. ISME J. 5, 1067–1071.
- Voget, S., Wemheuer, B., Brinkhoff, T., Vollmers, J., Dietrich, S., Giebel, H.A., Beardsley, C., Sardemann, C., Bakenhus, I., Billerbeck, S., Daniel, R., Simon, M., 2015. Adaptation of an abundant *Roseobacter* RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. ISME J. 9, 371–384.
- von Richthofen, F., 1886. Führer für Forschungsreisende. Berlin, Oppenheim, pp. 308–310.
- von Scheibner, M., Dörge, P., Biermann, A., Sommer, U., Hoppe, H.G. Jürgens, K., 2014. Impact of warming on phyto-bacterioplankton coupling and bacterial community composition in experimental mesocosms. Environ. Microbiol. 16, 718–733.
- Walker, C. B., de la Torre, J. R., Klotz, M. G., Urakawa, H., Pinel, N., Arp, D. J., Brochier-Armanet, C., Chain, P. S. G., Chan, P. P., Gollabgir, A., Hemp, J., Hügler, M., Karr, E., A., Könneke, M., Shin, M., Lawton, T. J., Lowe, T., Martens-Habbena, W., Sayavedra-Soto, L. A., Lang, D., Sievert, S. M., Rosenzweig, A. C., Manning, G., Stahl, D. A., 2010. *Nitrosopumilus maritimus* genome reveals unique mechanisms for

- nitrification and autotrophy in globally distributed marine Crenarchaea. Proc. Natl. Acad. Sci. USA 107, 8818-8823.
- West, N. J., Obernosterer, I., Zemb, O., Lebaron, P., 2008. Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. Environ. Microbiol. 10, 738–756.
- Yabuuchi, E., Kaneko, T., Yano, I., Moss, C., Miyoshi, N., 1983. *Sphingobacterium* gen. nov., *Sphingobacterium spiritivorum* comb. nov., *Sphingobacterium multivorum* comb. nov., *Sphingobacterium mizutae* sp. nov., and *Flavobacterium indologenes* sp. nov.: glucose-nonfermenting Gram-negative rods in CDC Groups I1K-2 and I1b. Int. J. Syst. Evol. Microbiol. 33, 580–598.
- Yamada, Y., Kaga, S., Ogata, T., 2012. Influence of a huge tsunami on the coastal plankton community structure, especially on the abundance of the toxic dinoflagellate (*Alexandrium tamarense*) in Ofunato Bay. In: Program and Abstracts of PICES-2012. p. 146
- Yamada, Y., Kaga, S., Kaga, Y., Naiki, K., Watanabe, S., 2017. Changes of seawater quality in Ofunato Bay, Iwate, after the 2011 off the Pacific coast of Tohoku Earthquake. J. Oceanogr. 73, 11-24.
- Zaidi, B.R., Hinkey, L.M., Rodríguez, N.R., Govind, N.S., Imam, S.H., 2003. Biodegradation of toxic chemicals in Guayanilla Bay, Puerto Rico. Mar. Pollut. Bull. 46, 418–423.

Legends to Figures

Fig. 1. Location of three seawater sampling stations in the Ofunato Bay, KSt. 1, 2 and 3, marked with arrows.

Fig. 2. GC% profiles of KSt. 1, 2 and 3WGS metagenomic datasets collected from 1 m depth (0.2 μ m size filter fraction only) of the Ofunato Bay.

Fig. 3. Phylogenetic profiles of the Ofunato Bay metagenome. Bacterial and archaeal profiles are shown using two different approaches, one using the 16S rRNA sequences against Ribosomal Database Project (16S) gathered from whole genome sequencing datasets, and the other using all the reads on shotgun metagenomics and comparing against NCBI-nt database (WGS).

Table 1

Metagenome properties obtained for three sampling stations in the Ofunato Bay.

Label	KSt. 1	KSt. 2	KSt. 3
Sequence size (bp)	341,161,985	313,609,427	277,686,572
Sequence size after filtering (bp)	340,685,026	313,342,919	277,237,384
No. of reads	1,225,081	1,021,226	934,064
No. of reads after filtering	1,213,968	1,015,007	923,558
Sequence length range after filtering (bp)	50 - 592	50 - 592	50 - 592
Mean sequence length after filtering (bp)	313	313	313

Table 2

16S rRNA metagenome statistics obtained for three sampling stations in the Ofunato Bay.

Label	KSt. 1	KSt. 2	KSt. 3
No. of total sequences affiliated to 16S rRNA genes	2,705	2,453	2,028
No. of bacterial 16S rRNA gene sequences	2,567	2,363	1,973
No. of archaeal 16S rRNA gene sequences	138	90	55

Figure 1

Figure 2

Figure 3