



Metagenomic Methods: From Seawater to the Database

1

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

Abstract

In this article, methods or techniques of metagenomics including targeted 16S/18S rRNA analyses and shotgun sequencing will be discussed. It is sometimes difficult, especially for beginners, to follow the manufacturer's recommendation as mentioned in the protocol and to go through different steps from the preparation of starting material (e.g., DNA), library preparation, and so on. We

Md. S. Reza

Kitasato University School of Marine Biosciences, Sagamihara, Japan

Department of Fisheries Technology, Bangladesh Agricultural University, Mymensingh, Bangladesh

A. Kobiyama · J. Rashid · Y. Yamada · Y. Ikeda · D. Ikeda · N. Mizusawa · S. Yanagisawa
S. Sato · T. Ogata · T. Kudo

Kitasato University School of Marine Biosciences, Sagamihara, Japan

K. Ikeo

Kitasato University School of Marine Biosciences, Sagamihara, Japan

National Institute of Genetics, Shizuoka, Japan

S. Kaga

Iwate Fisheries Technology Center, Iwate, Japan

Ofunato Fisheries Promotion Center, Iwate Prefectural Government, Iwate, Japan

S. Watanabe · K. Naiki · Y. Kaga · S. Segawa

Iwate Fisheries Technology Center, Iwate, Japan

K. Mineta · V. Bajic · T. Gojobori

Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

S. Watabe (✉)

Kitasato University School of Marine Biosciences, Sagamihara, Japan

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

© Springer Nature Singapore Pte Ltd. 2019

T. Gojobori et al. (eds.), *Marine Metagenomics*,

https://doi.org/10.1007/978-981-13-8134-8_1

3

will try to explain all the steps in detail and share our experience here. It all starts with collection of samples and collection of ecological/environmental metadata followed by sample fractionation (optional), extraction of DNA, sequencing, and finally data analyses to interpret results. Sample collection has always been the most important part of a study as it requires proper planning, a good workforce to execute, permission(s) of sampling from appropriate authority, and precaution(s) about endangered species during sampling. Here, we first describe methodology for a shallow river and in the later section methodology for a deep marine bay. In either case, slight modifications can be made to succeed in sampling. Determination of physicochemical parameters as metadata simultaneously is also an important task. These samples are then processed to extract DNA which needs to be representative of all cells present in the sample. Finally, sequencing is done by a next-generation sequencer, and data analyses are completed. Through these methods, scientists are now able to overcome the unculturability problem of more than 99% of environmental microorganisms and uncovered functional gene diversity of environmental microorganisms.

Keywords

Free-living microorganisms · Particle-associated microorganisms · Shotgun metagenomics · Size fractionation · Targeted metagenomics

1.1 Methods

Two approaches are used to study environmental samples including soil, water, and air:

1. Targeted 16S or 18S rRNA analysis
2. Shotgun sequencing

For either of these approaches, methodology of metagenomic workflow remains more or less similar, and they can be applied for any environment from shallow river to deep lake, sea or bay, air, and sediment.

1.1.1 Methodology for 16S rRNA Environmental Sequence Analyses

1.1.1.1 Sample Collection

The foremost part of environmental metagenomic study is sampling. Before planning for sampling, proper permission should be taken from the city government office, mayor's office, etc. Consent from fishery cooperatives may sometimes be necessary. In the case of sampling from deep waters, safety of the researchers is an important issue. Appropriate lifesaving gears should be taken. Care should be given to preserve endangered species or wildlife. Finally, proper cautions should be taken to ensure collection of representative samples as it may directly affect downstream analyses.



Fig. 1.1 Metagenomic sampling from a shallow river. (a) Monitoring environmental parameters, (b) recording environmental data, (c) sampling from surface layer of the river, (d) prefiltration using 100 μm filter, (e) pouring water sample into a sterile bottle, and (f) collection of sediments from riverbed of the Tama River, Tokyo, Japan

During collection of samples, proper care should be taken to minimize contamination (Fig. 1.1). It may be suggested to use sterilized glassware for collecting water samples from a river or sea. Alternatively, water collection vessels can be rinsed with water three times before being used for sample collection. In case of collection of sediments from the river or seabed, it is always better to use “Marine Biology Dredge,” e.g., Ekman Dredge, when the water is deep. Otherwise, sediments can be directly scrapped from the riverbed when the water depth is low.

1.1.1.2 Transport and Preservation of Sample

It is always important to quickly transport the samples to the laboratory for processing. Care should be taken not to allow rising of the temperature inside the samplers. It is suggested to use ice packs and coolers with an icebox for transportation, thereby minimizing multiplication of microorganisms inside the bottle. Sometimes, it may become necessary to preserve it as samples may be collected from places far from the laboratory. In case of lack of on-site sample preparation facilities, the samples can be preserved at 4 °C for 8–10 h.

1.1.1.3 Sample Treatment and Filtration

Size fractionation of water is an important part of metagenomic analyses as it separates planktonic microorganisms according to their size. It is a routine to use size fractionation, often equipped with a prefilter of larger pore size (typically 1.0–100 μm) and smaller pore sizes (0.1–0.2 μm) over a sample volume ranging from 0.05 to 20 l (Ganesh et al. 2014; Padilla et al. 2015; Reza et al. 2018a) (Fig. 1.2). It is generally accepted that free-living bacterioplankton that consist of more than 90% of the total bacterial population in the marine and freshwater ecosystems are trapped on the small-size pore filters, while their particle-associated counterparts get trapped on the larger-size pore filters as they remain attached to suspended particles particularly in the 5- μm filter fraction during filtration (Reza et al. 2018a). While the wide range of sample volumes used for filtration in different samplings is typical for coastal and open water environments (Padilla et al. 2015), the use of prefilter by larger pore size is generally practiced to remove debris that may be present in the sample (Reza et al. 2018b).

Different types of water filtration systems are available in the market such as:

- (i) Nalgene™ polysulfone reusable bottle top filter with receiver (Thermo Fisher Scientific, Waltham, MA, USA)
- (ii) Advantec purification system (Advantec Co. Ltd., Tokyo, Japan)
- (iii) Sterivex™ cartridge filters (Millipore, Merck KGaA, Darmstadt, Germany)

In the case of using Nalgene™ polysulfone reusable bottle top filter with a receiver, ~350 ml of water is poured onto the filter with receiver (Thermo Fisher Scientific) and filtered sequentially through 45-mm diameter each of 5-, 0.8-, and 0.2- μm pore-sized Isopore™ membrane filters (Merck Millipore Ltd., Tullagreen, Ireland) using a vacuum pump. The pressure of the vacuum pump should be monitored carefully as it may disrupt cells of many soft-bodied protists and eukaryotes.

1.1.1.4 DNA Extraction

Generally microorganisms adhere to the filters, and different methods are used to extract DNA from these filters. Here we will discuss Qiagen's bead-beating

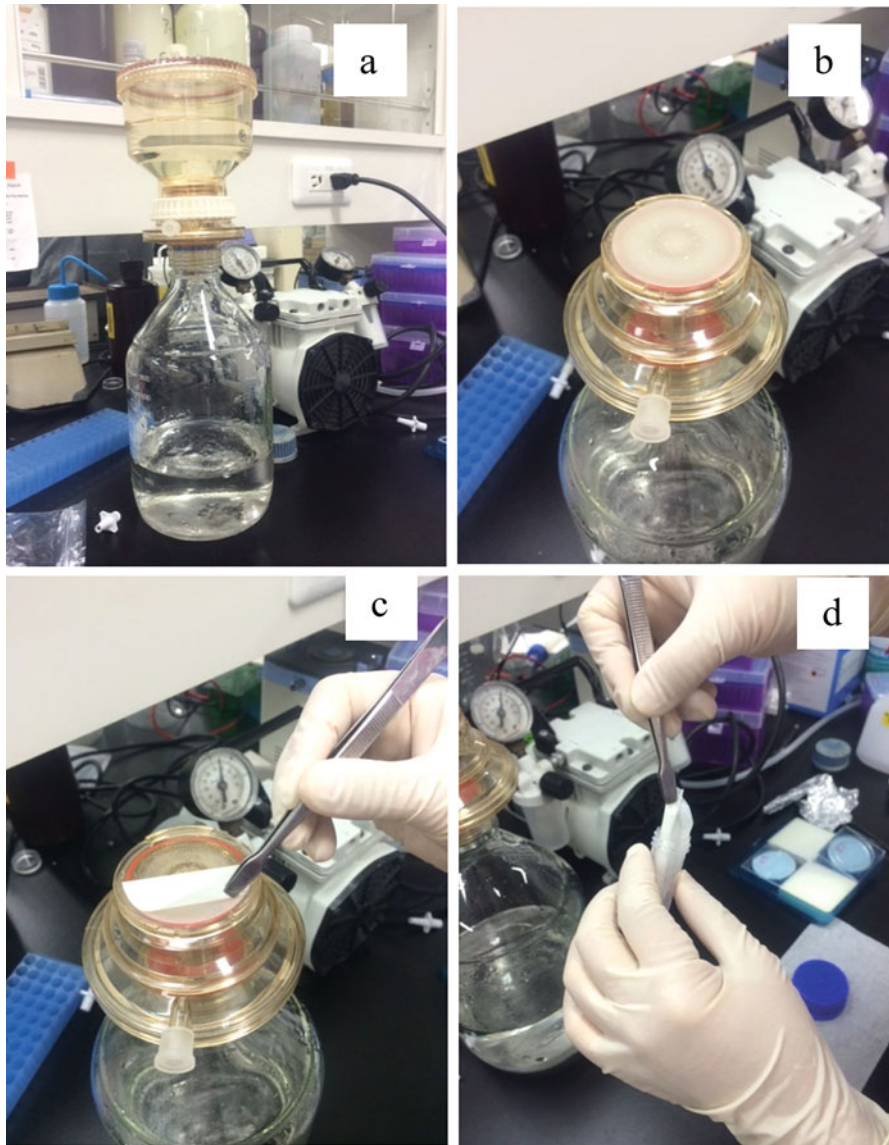


Fig. 1.2 Size fractionation of water samples. (a) Filtration of water, (b) trapping microorganisms on a filter, (c) rolling a filter paper, and (d) insertion of a filter paper into a 5-ml PowerWater Bead Tube

technique (Qiagen GmbH, Hilden, Germany). After removing the filters from the filtration unit using sterile forceps, the following steps can be used:

- (i) First the filter membrane should be inserted into the 5-ml PowerWater[®] Bead Tube and subsequently be rolled in such a way that its top side faces inward.
- (ii) Then 1 ml of Solution PW1 preheated at 55 °C is added to the PowerWater[®] Bead Tube.
- (iii) The PowerWater[®] Bead Tube is secured horizontally to a Vortex adapter and vortexed at the maximum speed for 5 min.
- (iv) Then the tubes are centrifuged at $\leq 4000 \times g$ for 1 min at room temperature. The speed will depend on the capability of the centrifuge in use.
- (v) The supernatant is transferred to a clean 2-ml Collection Tube by drawing up the supernatant using a 1-ml pipette tip by placing it down into the beads. Here, placing the pipette tip down into the beads and pipetting more than once are required to ensure the removal of the supernatant. Any carryover of beads will not affect subsequent steps. In this way, recovery between 600 and 650 μl of the supernatant is possible depending on the type of filter membranes used.
- (vi) Then the tubes are centrifuged at $13,000 \times g$ for 1 min.
- (vii) To avoid the pellet, the supernatant is transferred to a clean 2-ml Collection Tube, and 200 μl of Solution PW2 are added, and the mixture is vortexed briefly to mix. They are subsequently incubated at 4 °C for 5 min and centrifuged at $13,000 \times g$ for 1 min.
- (viii) Then the supernatant is transferred to a clean 2-ml Collection Tube by carefully avoiding the pellet.
- (ix) To the 2-ml Collection Tube, 650 μl of Solution PW3 is added, and the mixture is vortexed briefly to mix. Then 650 μl of supernatant is loaded onto a Spin Filter and centrifuged at $13,000 \times g$ for 1 min.
- (x) The flow is discarded through, and the process is repeated until all the supernatant has been loaded onto the Spin Filter. In this way, a total of two loads for each sample processed are required.
- (xi) Then the Spin Filter basket is placed into a clean 2-ml Collection Tube, and 650 μl of Solution PW4 is added and centrifuged at $13,000 \times g$ for 1 min.
- (xii) The flow through is discarded, and 650 μl of Solution PW5 is again added, and the mixture is centrifuged at $13,000 \times g$ for 1 min. The flow through is again discarded and centrifuged again at $13,000 \times g$ for 2 min to remove residual wash.
- (xiii) Finally the Spin Filter basket is placed into a clean 2-ml Collection Tube and centrifuged at $13,000 \times g$ for 1 min after adding 100 μl of Solution PW6 to the center of the white filter membrane.
- (xiv) The Spin Filter basket is discarded, and the obtained DNA is now ready for any downstream application.

After extraction, DNA concentration needs to be measured as most of the library preparation kits require a specified amount of DNA for downstream analyses. For

example, Nextera XT DNA Sample Preparation Kit (Illumina Inc., San Diego, CA, USA) requires only 1 ng of input DNA. In such cases, DNA quantitation can be done with a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) using Qubit dsDNA HS Assay Kit (Invitrogen) that uses highly sensitive and accurate fluorescence-based Qubit quantitation assays for quantitation of DNA, RNA, and protein. For each assay, it may be necessary to run new standards for calibrating the Qubit 2.0 Fluorometer. For later uses, the values from the previously done calibration can be used once calibration is completed for the first time. The process is briefly described below:

- (i) Calibration of the Qubit 2.0 Fluorometer requires the preparation of the appropriate standard solutions. We need to set up two assay tubes: one for standards and the other for our sample. To prepare the Qubit Working Solution by diluting the Qubit reagent 1:200 in Qubit buffer, we need 200 μl of working solution each for standard and sample. We need to add 190 μl of working solution and 10 μl of standard solution (supplied with the kit) to make total volume in each assay tube to 200 μl . As for assay of our sample, we need to add 180–199 μl working solution and 1–20 μl of our sample to make the total volume to 200 μl . Generally, we use 1–2 μl of our sample and 198–199 μl of working solution for each sample assay.
- (ii) Then all tubes are to be vortexed for 2–3 s and incubated at room temperature for 2 min.
- (iii) After completion, they should be inserted into the Qubit 2.0 Fluorometer to take readings.

The extracted DNA can be stored at $-20\text{ }^{\circ}\text{C}$, and aliquots of the extracted DNA can be used for metagenomic 16S rRNA sequencing.

1.1.1.5 PCR Amplification of 16S rRNA Gene and Next-Generation Sequencing

To determine the diversity and composition of the bacterial communities in the size-fractionated microorganism samples, different variable regions such as V1, V2, or V3 of the bacterial 16S rRNA gene can be targeted in metagenomic studies. The selection of appropriate primer is very important, as success of such study depends on the selection and design of primers. In this section, we discuss about sequencing methodology using Ion PGM (Thermo Fisher Scientific) and MiSeq (Illumina Inc., San Diego, CA) platforms.

- (i) For Ion PGM:

Polymerase chain reaction (PCR) is performed as follows: 20 μl each of PCR mixture comprises 0.1 μl Ex *Taq* HS DNA polymerase (Takara, Otsu, Japan), 2 μl of 10 \times Ex *Taq* DNA polymerase buffer (Takara), 1.6 μl of 2.5 mM dNTP, 12.3 μl of nuclease-free water, 1 μl of 2 mM adaptor-labeled forward primer 1 μl of 2 mM

adaptor-labeled and barcoded reverse primer, and 2 μ l of undiluted template DNA. PCR conditions can consist of 5 min incubation at 95 °C followed by 30 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s with the last cycle employing an extension for 5 min at 72 °C. Required number of independent PCRs can be carried out for the same amount of sample with a combination of forward primers and the common reverse primer.

Following amplification, all PCR products are examined for their size and specificity by electrophoresis on 2.5% w/v agarose and gel purified using Fast-Gene Gel/PCR Extraction Kit (Nippon Genetics Co. Ltd., Tokyo, Japan). Prior to sequencing, all amplicon types are assessed for fragment size distribution and DNA concentration using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). Ion Torrent sequencing is then performed according to the Ion Torrent standard workflow (Thermo Fisher Scientific). All barcoded amplicons are diluted to 13 pM and equimolarly pooled. Then 25 μ l of amplicon library is used for downstream application using the Ion PGM Template OT2 400 Kit for use with the Ion OneTouch 2 System. This procedure results in the generation of template-positive Ion PGM Template OT2 400 Ion Sphere Particles. After enrichment of the particles, a sequencing run is performed with the Ion 318 Chip Kit v2 based on the Ion PGM Sequencing 400 Kit User Guide, and FASTQ files are generated with Torrent Suite software version 4.2.1.

(ii) For MiSeq:

For targeted 16S rRNA analyses using MiSeq, PCR libraries need to be constructed using forward (F) and reverse (R) indexed primes targeting variable regions between V1 and V3. For making barcoded libraries using Illumina's TruSeq or Nextera kits, 12 barcodes come with the library preparation kit that Illumina provides. Therefore, indexing primers included in the sequencing reagent kits are used for running MiSeq (Reza et al. 2018c). After PCR amplification, the amplicons are quantified using a Bioanalyzer and pooled in equimolar amounts as described in Illumina's "16S Metagenomic Sequencing Library Preparation" Guide (Part # 15044223 Rev. B). Finally, sequencing is performed as paired-end reads using an Illumina MiSeq platform and a MiSeq Reagent Kit v3 (600 cycles).

1.1.1.6 Sequence Analyses

For Ion PGM datasets, individual sequence reads are first filtered within the PGM software to remove low-quality and polyclonal sequences. To avoid possible bias and to minimize taxonomic ambiguity, a stringent criterion is generally set to use only longer reads where sequences are again quality trimmed to remove <270 bp using Genomics Workbench 8.0 (CLC bio, Cambridge, MA, USA). They are then processed by the NGS analysis pipeline of the SILVA rRNA gene database project (SILVAngs 1.2) (Quast et al. 2013). Downstream analyses are performed as described by Klindworth et al. (2013), where different tools in the pipeline including SINA for the alignment of sequences (Pruesse et al. 2012), CD-HIT for

the clustering of sequences (Li and Godzik 2006), and BLAST for the classification of sequences (Camacho et al. 2009) are included.

As for 16S targeted libraries generated by MiSeq, all reads can be uploaded to BaseSpace server, and sequences of at least 100 bp are analyzed using the Illumina 16S Metagenomic App (v1.0.1) for taxonomic classification using an Illumina-curated version of Greengenes May 2013 reference taxonomy database.

1.1.1.7 Results Obtained for 16S rRNA Analyses for the Tama River, Tokyo, Japan

In this section, we present results of one of our studies conducted on Tama River, Tokyo, Japan, that employed next-generation sequencing technology targeting a 16S ribosomal RNA (rRNA) gene amplicon. Both the particle-associated and free-living portions collected from the river were studied separately after size fractionation of water samples trapped by 5.0-, 0.8-, and 0.2- μ m filters. It was revealed that this urban river was most dominated by Proteobacteria, followed by Bacteroidetes, Actinobacteria, and Cyanobacteria (Fig. 1.3). Through this approach, we did not detect any archaeal and eukaryotic sequences as we used bacteria-specific primers (Reza et al. 2018c).

Our study was the first river metagenomic report from Japan. Through the brief analysis, we were able to identify genus-level taxonomic profile of the Tama River that can work as a baseline for the river system concerned and can be used to

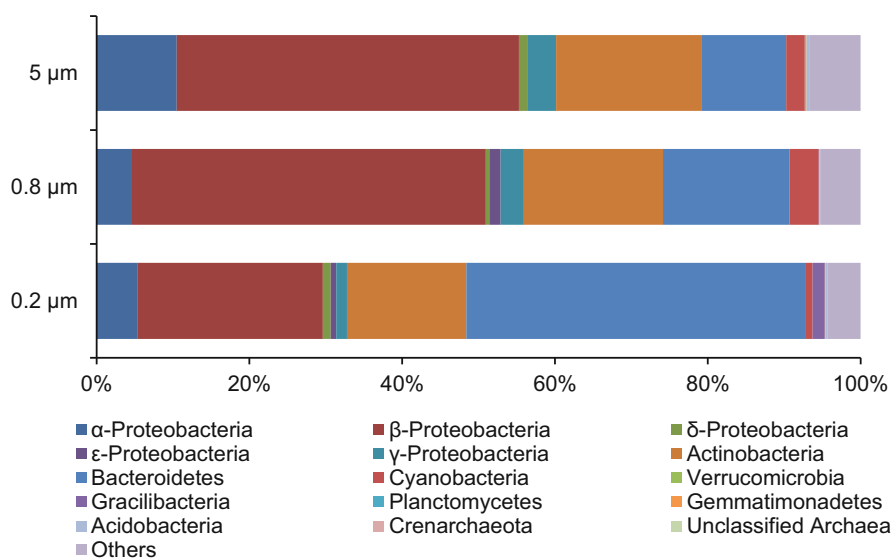


Fig. 1.3 Relative abundances of the major bacterial and archaeal phyla and classes in 5.0-, 0.8-, and 0.2- μ m filter fractions from the Tama River based on the taxonomic identification of 16S rRNA gene fragments. The artificial group “Others” contains unidentified sequences, as well as sequences with a relative abundance lower than 1%. (From Reza et al. 2018c)

enhance the existing water quality monitoring efforts to improve well-being of the residents in the area.

1.1.2 Methodology for Shotgun Metagenomic Sequence Analyses

1.1.2.1 Sample Collection

Generally for deep water expedition in the marine water, no specific permits are required for sampling. This gives freedom for researchers who conduct sampling. However, proper permissions should be obtained for protecting endangered or protected species at sea. Usually vertical water sampler is used to collect water sample from a specific water depth. The volume of water sample collected varies depending on biomass present in the water. Generally 5–10 l of seawater is collected and prefiltered through a 100- μm filter to remove debris (Fig. 1.4). Corresponding seawater quality parameters, including temperature, salinity, and dissolved oxygen, are also measured on site using a water quality profiler. Simultaneously representative water samples are collected in separate sunlight-protected bottles for measurement of other water quality parameters including the abundance of chlorophyll *a* (chl-*a*) and heterotrophic bacteria (Yamada et al. 2017).

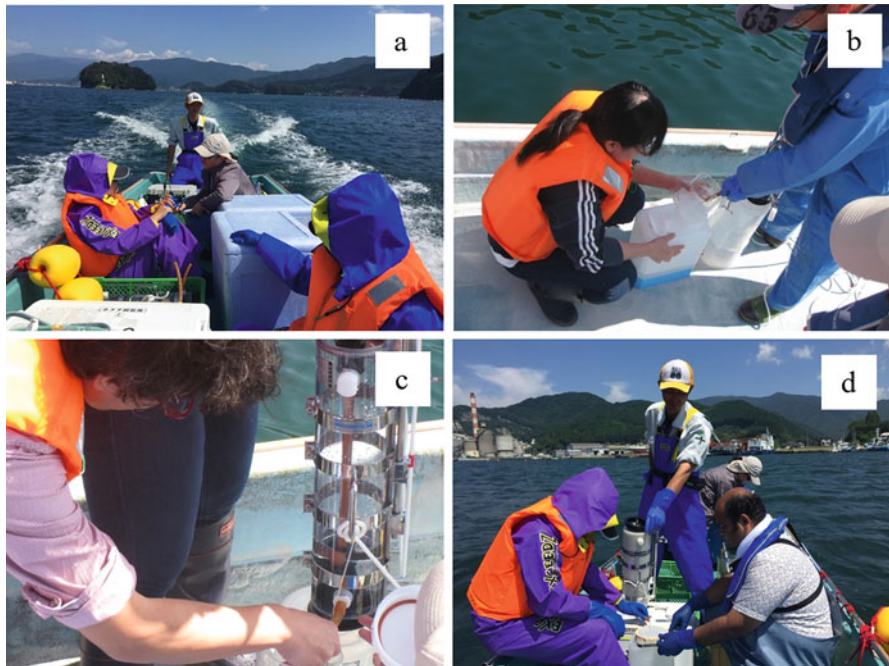


Fig. 1.4 Sampling of water at sea. (a) Travelling to sampling spot, (b) prefiltration using a 100- μm filter, and (c) and (d) sample collection in the Ofunato Bay, Iwate Prefecture, Japan

1.1.2.2 Sample Treatment and Filtration

In this section, we will discuss about filtration using Advantec purification system (Advantec Co. Ltd.). Generally the prefiltered water is filtered sequentially through 20- μm pore-sized Nylon Net Filter NY20 with a 47-mm diameter (Merck Millipore Ltd.) and through 5-, 0.8-, and 0.2- μm pore-sized Isopore Membrane Filters, 142 mm in diameter (Merck Millipore Ltd.), using a peristaltic pump (Masterflex L/S, Cole-Parmer International, Vernon Hills, IL, USA) (Fig. 1.5).

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 $-80\text{ }^{\circ}\text{C}$ until DNA extraction.

1.1.2.3 DNA Extraction

Genomic DNA is extracted from 0.2- μm filters using the PowerWater[®] DNA Isolation Kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA), following the manufacturer's recommendations as described earlier.

DNA concentrations also should be measured with Qubit dsDNA HS Assay Kit (Invitrogen), using a Qubit Fluorometer 2.0 (Invitrogen). The extracted DNA samples can be stored at $-20\text{ }^{\circ}\text{C}$ until use.



Fig. 1.5 Filtration of water. (a) Suction of water from tank, (b) and (c) sequential filtration using filtration system, and (d) peristaltic pump in use

1.1.2.4 Sequence processing and analyses

Extracted DNA aliquots are used for metagenomic shotgun sequencing. For library preparation by Nextera XT DNA Sample Preparation Kit (Illumina Inc.), 1 ng of gDNA is required. So sufficient amount of DNA should be diluted with sterile distilled water to adjust the concentration to 0.2 ng/ μ l.

Five microliters of gDNA at a concentration of 0.2 ng/ μ l are then taken, whereas DNA samples for metagenomics are prepared employing the Nextera XT DNA Sample Preparation Kit (Illumina Inc.) and sequenced as paired-end reads using an Illumina MiSeq platform and a MiSeq Reagent Kit v3 (600 cycles) (Illumina Inc.).

Illumina paired-end reads for each library are 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 base pairs and the overlap length, 0.25). Quality filtering of these shotgun sequencing reads is then performed by removing reads <50 bp and quality trimmed to Phred 20 using the Genomics Workbench (CLC bio). BLASTn is performed using the quality- and size-filtered sequences against the NCBI-nt reference database. Taxonomic analysis at the genus level is then performed using MEGAN v5.10.3 (Huson et al. 2007), after parsing the BLAST output. Comparative analysis in MEGAN is also performed after normalizing counts per the recommendations of the authors. The least common ancestor (LCA) assignment algorithm is set at the parameters: minimum support = 1; minimum score = 50; top percent = 10; and maximum expected 1.0E – 10.

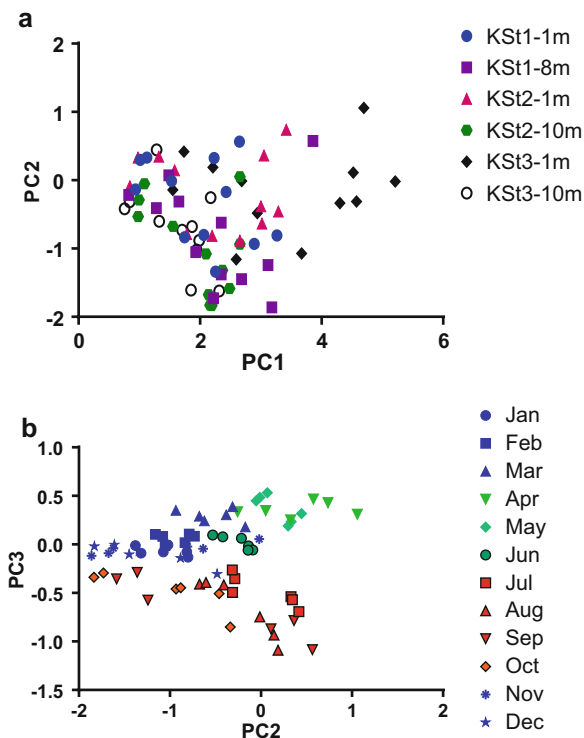
The obtained sequence matrices are generally used for correlating with various environmental factors including chl-*a* concentrations, heterotrophic bacterial abundances, and other metadata using statistical packages, e.g., R software in the Ofunato Bay (R Development Core Team 2008), based on experiment objectives (Rashid et al. 2018).

1.1.2.5 Results Obtained for a Shotgun Sequence Analyses for the Ofunato Bay, Iwate Prefecture, Japan

In this section, we are going to present the results of one of our shotgun metagenomic studies conducted on the Ofunato Bay, a deep enclosed bay in Iwate Prefecture, Japan (Reza et al. 2018b). Unlike our Tama River study, only free-living portions of the bay were reported as this portion has been considered as the most dominant among the total bacterial population. Phylogenetic differences among spring, summer, autumn, and winter bacterioplankton of the bay was studied that showed members of some group decreased at high water temperatures but increased at low temperatures, indicating seasonal change of the free-living bacterioplankton in the Ofunato Bay (Fig. 1.6).

As the Ofunato Bay is housed by Japan's famous buoy-and-rope type oyster (*Crassostrea gigas*) culture facilities, we were able to find strong locality signal for bacterial communities based on bacterial functional analyses (Kobiyama et al.

Fig. 1.6 Principal component analyses (PCAs) based on relative abundances of bacterial genera obtained from shotgun metagenomic data. Panel “a” shows that several datasets from KSt. 3 derived from 1-m water depth were clustered (dotted parabola), while Panel “b” shows that datasets derived from spring, summer, and winter were clustered together (parabola filled with green, red, and blue, respectively). Data employed was only from 0.2- μ m-size filters targeting the free-living fractions from samples collected during January to December in the Ofunato Bay. (From Reza et al. 2018b)



2018). The dominant bacterial species *Candidatus Pelagibacter ubique* and *Planktomarina temperata* were reportedly found to show change seasonally, playing important roles in producing dimethyl sulfide (DMS) and methanethiol (MeSH) from dimethylsulfoniopropionate (DMSP) as signaling molecules for possible formation of scent of the tidewater or as fish attractants (Kudo et al. 2018).

References

- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinform* 10:421–429
- Ganesh S, Parris DJ, DeLong EF, Stewart FJ (2014) Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J* 8:187–211
- Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data. *Genome Res* 17:377–386
- Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing based diversity studies. *Nucleic Acids Res* 41:e1
- Kobiyama A, Ikeo K, Reza MS, Rashid J, Yamada Y, Ikeda Y, Ikeda D, Mizusawa N, Sato S, Ogata T, Kudo T, Kaga S, Watanabe S, Naiki K, Kaga Y, Mineta K, Bajic V, Gojobori T, Watabe S (2018) Metagenome-based diversity analyses suggest a strong locality signal for bacterial communities associated with oyster aquaculture farms in Ofunato Bay. *Gene* 665:149–154

- Kudo T, Kobiyama A, Rashid J, Reza MS, Yamada Y, Ikeda Y, Ikeda D, Mizusawa N, Ikeo K, Sato S, Ogata T, Jimbo M, Kaga S, Watanabe S, Naiki K, Kaga Y, Segawa S, Mineta K, Bajic V, Gojobori T, Watabe S (2018) Seasonal changes in the abundance of bacterial genes related to dimethylsulfoniopropionate catabolism in seawater from Ofunato Bay as revealed by metagenomic analysis. *Gene* 665:174–184
- Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658–1659
- Magoc T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27:2957–2963
- Padilla CC, Ganesh S, Gantt S, Huhman A, Parris DJ, Sarode N, Stewart FJ (2015) Standard filtration practices may significantly distort planktonic microbial diversity estimates. *Front Microbiol* 6:547
- Pruesse E, Peplies J, Glöckner FO (2012) SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596. (database issue)
- Rashid J, Kobiyama A, Reza MS, Yamada Y, Ikeda Y, Ikeda D, Mizusawa N, Ikeo K, Sato S, Ogata T, Kudo T, Kaga S, Watanabe S, Naiki K, Kaga Y, Mineta K, Bajic V, Gojobori T, Watabe S (2018) Seasonal changes in the communities of photosynthetic picoeukaryotes in Ofunato Bay as revealed by shotgun metagenomic sequencing. *Gene* 665:127–132
- Reza MS, Kobiyama A, Yamada Y, Ikeda Y, Ikeda D, Mizusawa N, Ikeo K, Sato S, Ogata T, Jimbo M, Kudo T, Kaga S, Watanabe S, Naiki K, Kaga Y, Mineta K, Bajic V, Gojobori T, Watabe S (2018a) Taxonomic profiles in metagenomic analyses of free-living microbial communities in the Ofunato Bay. *Gene* 665:192–200
- Reza MS, Kobiyama A, Yamada Y, Ikeda Y, Ikeda D, Mizusawa N, Ikeo K, Sato S, Ogata T, Jimbo M, Kudo T, Kaga S, Watanabe S, Naiki K, Kaga Y, Mineta K, Bajic V, Gojobori T, Watabe S (2018b) Basin-scale seasonal changes in marine free-living bacterioplankton community in the Ofunato Bay. *Gene* 665:185–191
- Reza MS, Mizusawa N, Kumano A, Oikawa C, Ouchi D, Kobiyama A, Yamada Y, Ikeda Y, Ikeda D, Ikeo K, Sato S, Ogata T, Kudo T, Jimbo M, Yasumoto K, Yoshitake K, Watabe S (2018c) Metagenomic analysis using 16S ribosomal RNA genes of a bacterial community in an urban stream, the Tama River, Tokyo. *Fish Sci* 84:563–577
- R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0
- 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