We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Microbial Biodiversity and Biogeography on the Deep Seafloor

Shingo Kato and Akihiko Yamagishi
Department of Molecular Biology, Tokyo University of Pharmacy and Life Science,
Hachioji, Tokyo,
Japan

1. Introduction

Microbes are widely distributed on and within the Earth (Gold, 1992; Whitman et al., 1998). They have co-evolved with the Earth through their history and have acquired their diversity. Although most of microbes (estimated more than 99% of the total species numbers present on Earth) are still uncultivated (Amann et al., 1995), vigorous surveys in natural environments, from cold poles to hot deep-sea vents, by microbiologists revealed the wide distribution of microbes. The accumulation of 16S rRNA gene sequence data and the development of useful bioinformatics tools allow us to image a big picture of microbial biodiversity and biogeography in natural environments (Martiny et al., 2006). This will help us to address some fundamental questions about microbial community: where they are; what species are present there; how they constitute communities; what are the factors that control the diversity and distribution pattern of the communities; and how they have evolved from the past to present and will evolve in future. Considering the wide distribution and powerful metabolic function of microbes, they are likely to contribute to the maintenance of the whole ecosystem on Earth and to global climate change.

1.1 Biodiversity and biogeography

"Microbial biodiversity" (or simply "biodiversity", "microbial diversity") includes phylogenetic (genotypic) and physiologic (phenotypic) diversity of microbial communities. In many previous papers (and also in this chapter), microbial biodiversity indicated the phylogenetic diversity that can be measured based on variation in nucleotide sequences of genes (16S rRNA gene have been used frequently). The biodiversity measures can be distinguished into the measure of diversity within a single community (α-diversity) and that of the partitioning of diversity among two or more communities (β-diversity). The biodiversity can be measured qualitatively (based on the presence/absence of each taxon) and quantitatively (taken account for the abundance of each taxon). Furthermore, species-based (treated all taxa as equally) and divergence-based (taken account for the phylogenetic distance between each taxon) measurements are used. The above classification and methods for biodiversity measurements have been well summarized (Lozupone & Knight, 2008). To avoid misinterpretation of results, it is needed to understand the principles of these diverse measurements for microbial biodiversity. Because the microbial diversity potentially affects

on the ecosystem functioning (Duffy & Stachowicz, 2006; Prosser et al., 2007), it is important to measure and interpret correctly the microbial diversity in natural environments.

Microbial biogeography is the descriptive and explanatory study of microbial biodiversity over space and time. It aims to reveal where microbes live, and what kinds of microbes are present there and how many they exist. The scope of biogeography extends to understand the underlying mechanism of generating and maintaining the distribution pattern of microbial communities in natural environments. Traditional biogeography has focused on large eukaryotes, such as plants and animals. Recent development of molecular biological techniques enabled us to approach the biogeography for microbes, including protists and prokaryotes in gene sequence level (Darling et al., 2000; Whitaker et al., 2003). This offers a challenge to the famous classical Baas Becking hypothesis *'Everything is everywhere*, but, *the environment selects'* (Baas Becking, 1934). To assess how the environmental similarity (i.e., contemporary physicochemical condition) and geographic distance (i.e., historical event) affect on the distribution patterns of biodiversity of microbial communities or population, β-diversity are associated with physicochemical and geographic differences among each environment. A variety of habitats in natural environments have been targeted for microbial biogeography, such as soil (Cho & Tiedje, 2000; Fierer & Jackson, 2006), lake sediments (Yannarell & Triplett, 2005), ocean (García-Martínez & Rodríguez-Valera, 2000), deep-sea sediments (Schauer et al., 2009) and deep-sea hydrothermal fields (Kato et al., 2010). These studies have suggested that the distribution pattern of microbes in habitats is controlled by not only environmental factors (e.g., pH, salinity and oxygen concentration) but also geographic isolation. However, further data collection and reliable explanation are needed to propose and evaluate theories regarding the generation and evolution of distribution pattern of microbes in natural environments. While the study of microbial biodiversity and biogeography seems to be descriptive, it is the first step and the essential base for theory construction in microbial ecology (Prosser et al., 2007).

1.2 Seafloor microbial communities

Deep seafloor is seemingly an unrelieved, monotonous and poor environment, like a desert where organisms are scarcely present. Actually, this notion is not always correct. There are variable environments on the deep seafloor, such as hydrothermal vents, cold seeps, iron-rich mats, out crops of young crustal rocks and aged ferromanganese crusts (hereafter, Mn crusts). Furthermore, phylogenetically and physiologically diverse microbes (especially prokaryotes, i.e., the domain *Bacteria* and *Archaea*) thrive in these environments (e.g., Takai & Horikoshi, 1999; Inagaki et al., 2002; Santelli et al., 2008; Kato et al., 2009a; Nitahara et al., 2011). Following the Baas Becking hypothesis, these microbes may adapt to each environment and should form a unique community structure. However, the hypothesis has not been tested well for the microbial communities on the deep seafloor, especially non-hydrothermal and unsedimented areas far from land where organic inputs derived from surface photosynthetic ecosystems are not significant. Such deep seafloor accounts for a large part of the surface area of Earth (Smith & Sandwell, 1997). Microbes on and within the seafloor are thought to play a role in geochemical cycling between oceans and Earth crusts (Edwards et al., 2005). Hence, understanding the microbial diversity and biogeography on the deep seafloor is important for modeling the global relationship between microbes and Earth at present and can be applied for in past and future. Furthermore, recently, massive sulfide deposits and Mn crusts on the deep seafloor have been focused on as mineral

resources (Rona, 2003; Hoagland et al., 2010). There are diverse microbes on/in these seafloor minerals (Kato et al., 2010; Nitahara et al., 2011). Microbial biogeography on the deep seafloor will contribute ultimately to develop deep-sea mining techniques utilizing microbes in future.

2. Analytical methods

The study of microbial biodiversity and biogeography starts upon collecting data of microbial communities in the environments. In this chapter, we introduce the analysis methods of microbial communities based on nucleotide sequences of genes, especially the small subunit ribosomal RNA gene (called 16S rRNA gene for prokaryotes) which is generally used for such analysis of biodiversity and biogeography. 16S rRNA genes have some merit for the analysis: 1) all prokaryotes have this gene; 2) its sequence length (approximately 1500 bases) is moderate and adequate for analysis; 3) there are some conserved regions that allow to design PCR primers; 4) there are some variable regions that allow to affiliate the sequences in species-level; 5) enormous sequences have been deposited in public database and can be used conveniently; and 6) useful bioinformatics tools specialized for this gene are available.

Recent rapid development of molecular biological techniques including gene amplification and nucleotide sequencing enabled us to approach unexpected biodiversity of microbes in natural environments. Especially, next-generation sequencing techniques (e.g., pyrosequencing) open new windows for approaching microbial biodiversity (Sogin et al., 2006). However, it is hard to gain biological meanings of the biodiversity resulted from short sequences (~400 bases) that are produced by the next-generation sequencing. At least, nearly full-length sequences of 16S rRNA genes are needed to connect biodiversity to ecology. It should be noted that information on the gene sequences cannot be related to ecology directly. To connect biodiversity to ecology, the determination of the whole-genome sequence is not enough and information on function of microbes derived from the gene sequences must be obtained by culture-dependent analysis. However, cultivation of all microbes in an environment is impossible by now. Actually, most prokaryotes on Earth are still uncultivated (Amann et al., 1995).

There are several steps to measure and compare the biodiversity of microbial communities based on 16S rRNA gene analysis (Figure 1). The target environmental samples are collected, genomic DNA is extracted from the samples, and then 16S rRNA gene sequences in the genome DNA extracts are determined by PCR-cloning-sequencing analysis. In some cases, analyses of electrophoresis patterns, such as denaturing gradient gel electrophoresis (DGGE) and terminal restriction fragment length polymorphism (T-RFLP), without sequencing process are also used for biodiversity measurement; however these analyses mask valuable information on the biodiversity in contrast to sequencing analysis (Nocker et al., 2007). It should be noted that PCR-cloning-sequencing analysis alone is insufficient for the determination of biodiversity because of the presence of methodological biases (Wintzingerode et al., 1997) in addition to relatively high cost performance regarding time and money. For example, even if the same DNA extract was used, the diversity and composition of microbial communities determined using different primer sets were dramatically different from each other (Kato et al., 2011). Hence, for reliable assessment of the biodiversity and distribution pattern of microbial communities, 16S rRNA gene sequences used for comparative analysis should be obtained by the same method.
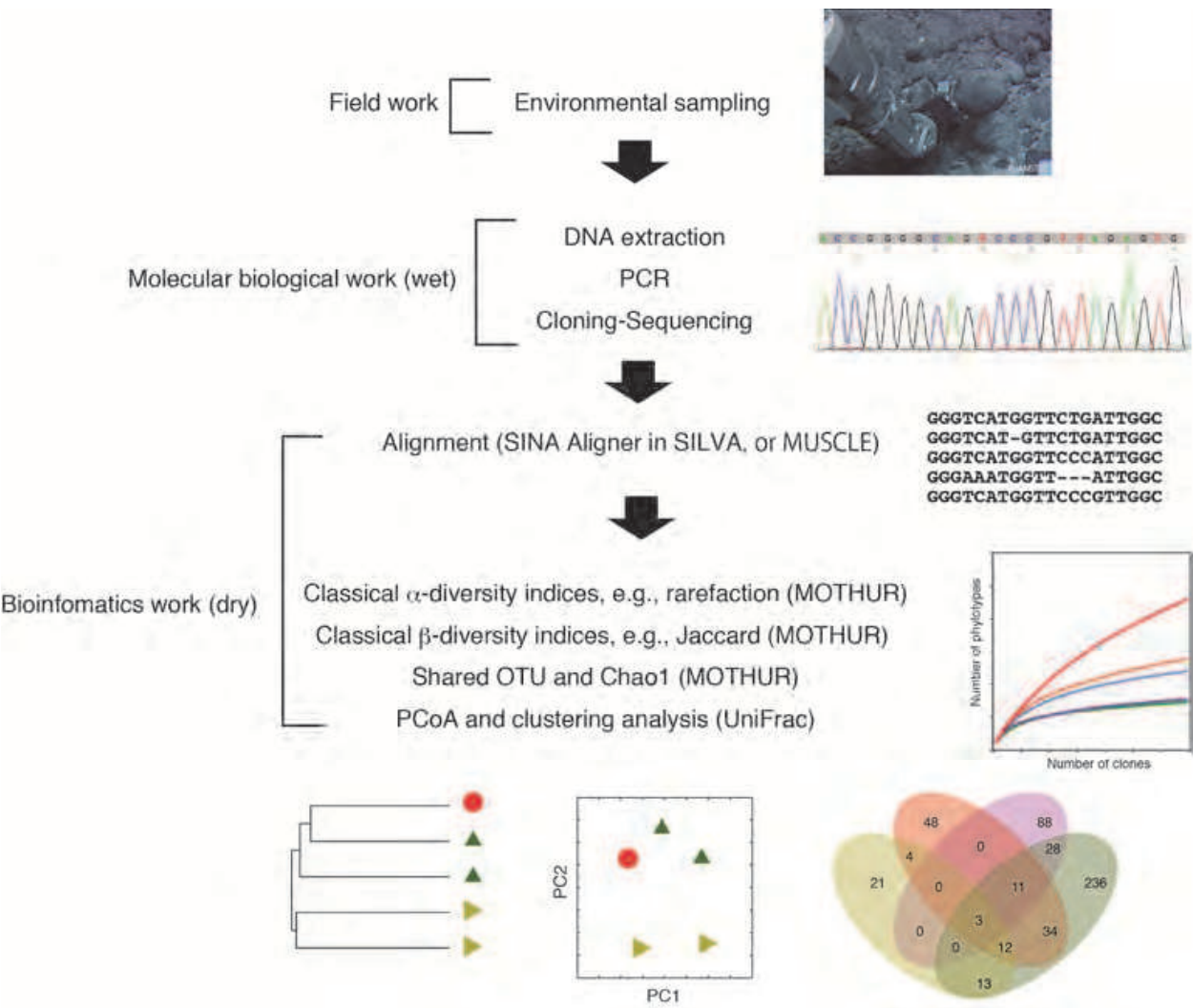
Fig. 1. Flow chart of the recommended analysis steps for microbial diversity and biogeography.

All of the 16S rRNA gene sequences collected are aligned into an alignment dataset. The alignment process is very important to assess more accurate biodiversity and biogeography because the wrong alignment dataset cause overestimation of biodiversity and of difference among communities. Alignments have often been performed by multiple sequence alignment tools such as ClustalW (Larkin et al., 2007) and MUSCLE (Edgar, 2004). However, against the vast 16S rRNA gene datasets including one thousand sequences or more, it takes an immense amount of time. Recently, improved alignment methods incorporating the secondary structure of 16S rRNA genes and using a reference alignment have been provided from several 16S rRNA gene database projects, such as RDP using Infernal (Cole et al., 2009), Greengenes using NAST aligner (Desantis et al., 2006), and SILVA using SINA aligner (Pruesse et al., 2007). However, these methods, in particular NAST, have predicted higher diversity as compared with the results from the pair-wise and multiple alignment methods (Schloss, 2010). The sequences with long insertions seem to be better aligned by Infernal built in RDP than the other methods from our experience. It is known that several prokaryotes, especially *Archaea*, have long insertions (including introns) in their 16S rRNA gene (Burggraf et al., 1993; Itoh et al., 1998; Itoh et al., 2003). Chimera sequences have often

observed in datasets. Such chimera sequences must be removed from the datasets before the following analysis by using chimera-check tools such as Mallard (Ashelford et al., 2006) and Bellerophon (Huber et al., 2004). NAST automatically checks chimera sequences in the datasets and remove the chimeric part of the sequences. However, the sequence with relatively longer insertion seems to be also recognized as a chimera sequence by NAST. Until more improvement of the alignment method and chimera check function, we recommend not using NAST for the following phylogenetic analysis. Overall, for alignment of 16S rRNA gene sequences, SINA aligner built in SILVA, or MUSCLE in the case of lower sequence numbers, is recommended for alignment. Finally, accuracy of the alignment dataset should be confirmed by the naked eyes.

After the construction of a 16S rRNA gene alignment dataset, the sequences are assigned as operational taxonomic units (OTUs) or phylotypes for each habitat. An OTU is a group of similar sequences each other, which is defined based on the genetic distance thresholds. In general, 97% (0.03 cut-off), 95% (0.05 cut-off) or 80% (0.20 cut-off) similarity threshold are used as species-, genus- and family-level taxonomic definition, respectively (Ludwig et al., 1998). For comparative analysis of communities, the same definition level of OTUs must be used. Assessment of sequences to OTUs can be performed using DOTUR (Schloss & Handelsman, 2005) and its current version mothur (Schloss et al., 2009). A distance matrix generated from the alignment dataset using ARB (Ludwig et al., 2004) or DNADIST in PHYLIP package (Felsenstein, 1989) is needed for calculation using DOTUR. The matrix can be generated by mothur itself.

α-diversity measures can be calculated using the distance matrix generated from the alignment dataset: Chao1 species richness estimates, abundance-based coverage estimator of species richness (ACE) and rarefaction curves (species-based, qualitative), Shannon and Simpson indices (species-based, quantitative), Phylogenetic Diversity (PD; divergence-based, qualitative) and θ (divergence-based, quantitative), and so on (Lozupone & Knight, 2008). Species-based measurements of α-diversity can be performed using mothur at once. PD and θ can be calculated by PHYLOCOM (Webb et al., 2008) and ARLEQUIN (Excoffier et al., 2005), respectively.

β-diversity can be also measured using the distance matrix generated from the alignment dataset: Sørensen and Jaccard indices (species-based), UniFrac and F_{ST} (divergence-based), and so on (Lozupone & Knight, 2008). These β-diversity values provide measures of distance between pairs of communities. Furthermore, the measured distance matrix can be used for multivariate statistical techniques such as clustering [e.g., unweighted pair group method using arithmetic average (UPGMA)] and ordination [e.g., principal coordinate analysis (PCoA)]. Several species-based measures of β-diversity can be calculated using SONS (Schloss & Handelsman, 2006), which has been incorporated into mothur. UniFrac (Lozupone & Knight, 2005; Lozupone et al., 2006), and its current version Fast UniFrac (Hamady et al., 2009), is an effective divergence-based method for β-diversity (Lozupone et al., 2010) and can easily perform clustering and ordination analyses. For UniFrac analysis, a phylogenetic tree and definition data for OTUs and habitats are needed as input data. This tree can be constructed from the alignment dataset by neighbor-joining (NJ) or maximum-likelihood (ML) method. NJ tree can be constructed using ARB or ClearCut (Sheneman et al., 2006). ML tree can be constructed using FASTTREE (Price et al., 2010) or PHYML (Guindon et al., 2010). Such clustering and ordination analyses can also be performed using R (R Development Core Team, 2011). In addition, the shared OTU numbers and the shared

Chao1 richness among communities can be calculated and viewed automatically in Venn diagrams using mothur. Overall, our recommendation of the analytical steps is shown in Figure 1.

3. Microbial biodiversity and biogeography of microbial communities on deep seafloor

3.1 Data collection

In this chapter, we try to assess microbial biodiversity and biogeography on deep seafloor using the recent useful bioinformatics tools as described above, though few data from non-hydrothermal and unsedimented deep-seafloor in open sea are available for biogeographical analysis. To investigate the distribution pattern of microbial communities on the seafloor in open sea, several data were collected (Table 1) and used for the following analysis. The locations where the samples were collected are shown in Figure 2. The samples of the collected data are basaltic rocks, Mn crusts, sulfide deposits called as dead chimney which were collected from hydrothermally inactive vents, sandy sediments that were not organic-rich, and overlying bottom seawater (Figure 2). Although the samples were mainly collected on spreading ridges, they were collected far from hydrothermal vents and may not be significantly influenced by hydrothermal activity. We analyzed and compared the communities as described above using the 16S rRNA gene sequences collected.

3.2 α-diversity

α-diversity for a microbial community is often indicated using Chao1 species richness estimates, rarefaction curves and/or Shannon's index value. However, we should not simply compare these indicators of α-diversity provided by the investigators because these indicators are biased by the PCR primers, alignment software and OTU or phylotype clustering methods used in the analyses. As is often the case, the sequences deposited into public databases do not contain all clones in the libraries but only the representative OTUs. Furthermore, the definition levels of OTUs are not always consistent; $OTU_{0.03}$ (i.e., 97% similarity level) are usually used, but $OTU_{0.01}$ or others are also used in some cases. In general, $OTU_{0.03}$ or $OTU_{0.05}$ is used as species or genus level definition, respectively. In this chapter, to compare α-diversity for the communities for the collected data as impartially as possible, the percentage of the number of $OTU_{0.05}$ in total clone numbers, $N_{0.05}/N_t$, were used (Table 1). This comparison can roughly address the difference in the α-diversity, even if only representative sequences were deposited and several definition levels of OTU (<0.05 cut-off) were used.

The $N_{0.05}/N_t$ are summarized in Figure 3. The $N_{0.05}/N_t$ for the samples of dead chimneys and seawater, except Asp, were <40%. In contrast, those for the samples of Mn crusts, basaltic rocks, except Rh3, and sediments were 50% or higher. Noted that the $N_{0.05}/N_t$ for Re5 and Rj were very high (>90%) likely due to the small total clone numbers compared with the other rock or sediment samples (Table 1). The relatively high $N_{0.05}/N_t$ of Asp may be due to contamination from seafloor sediment at the sampling. In fact, some phylotypes recovered from Asp were closely related to those from seafloor rocks and sediments (Kato et al., 2009b). The $N_{0.05}/N_t$ did not correlate with the total clone numbers analyzed ($r^2 = 0.223$) when the data of Asp, Re5 and Rj were excluded (the plot is not shown). Thus, the difference in the α-diversity associated with each habitat type is potentially meaningful for microbial
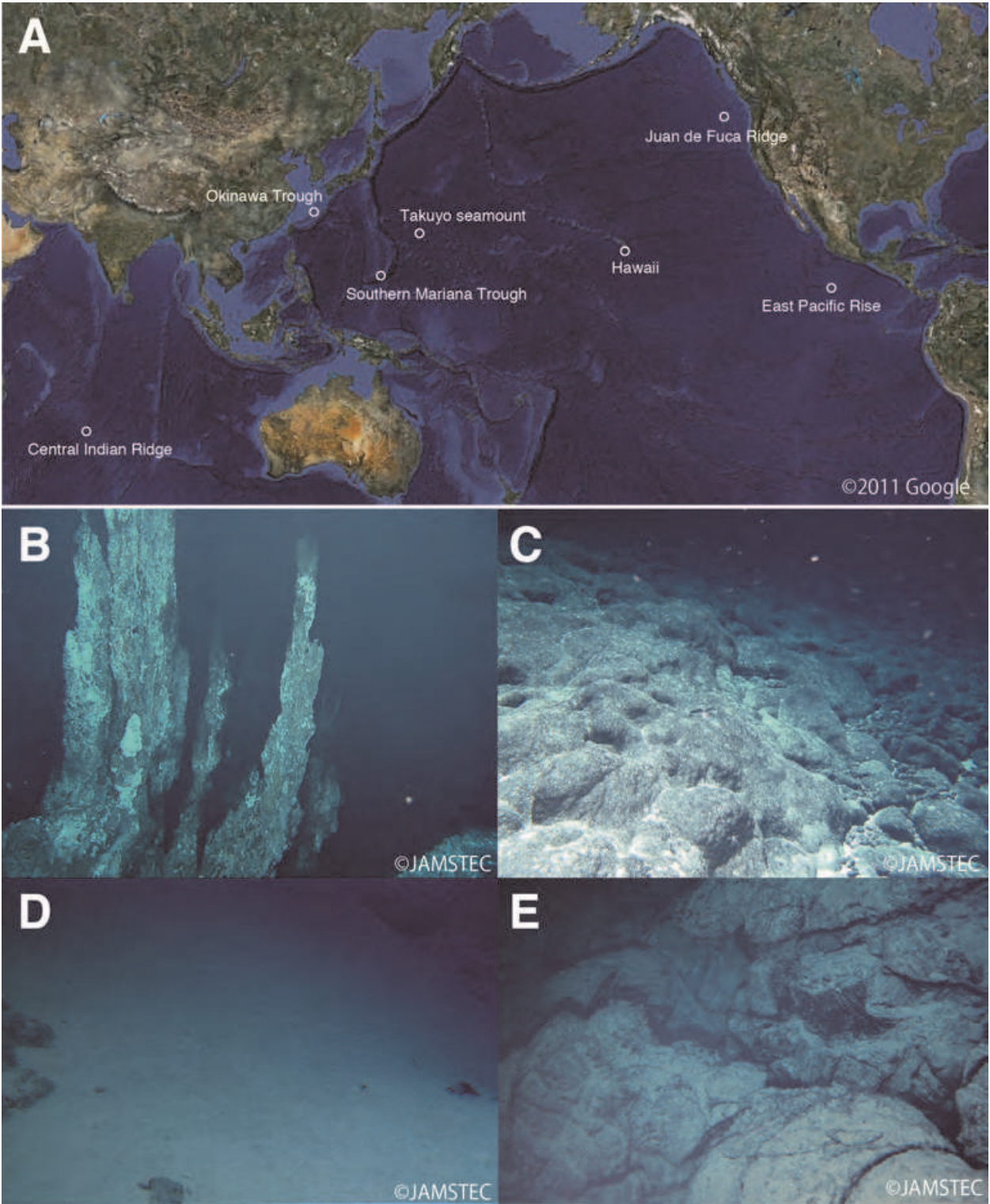
Fig. 2. (A) Location of the samples of which data were used in this chapter, and photographs of on-site observation of (B) sulfide chimney, (C) Mn crust, (D) sandy sediment and (E) basaltic rocks. These photographs (B-E) were taken at the Southern Mariana Trough and Takuyo-Daigo Seamount.

ID	Sample type	Location	Longitude	Latitude	Library name	Total clone number	$OTU_{0.05}$ number	$OTU_{0.05}$ number /Total clone number	Ref
Ae	Ambient seawater	East Pacific Rise	9°28.89'N	104°14.78'W	4055-N3	62	11	17.7	(Santelli et al., 2008)
Asf	Ambient seawater	Southern Mariana Trough	12°57.22'N	143°37.16'E	Fasw	45	17	37.8	(Kato et al., 2009b)
Asp	Ambient seawater	Southern Mariana Trough	12°55.15'N	143°36.96'E	Pasw	44	28	63.6	(Kato et al., 2009b)
At	Ambient seawater	Takuyo seamount	22°40.98'N	153°14.62'E	953sw	100	37	37.0	(Nitahara et al., in press)
Dc	Dead chimney	Central Indian Ridge	25°19.17'S	70°02.43'E	Ind	242	18	7.4	(Suzuki et al., 2004)
Do	Dead chimney	Okinawa Trough	27°47.48'N	126°53.78'E	Ihe	174	10	5.7	(Suzuki et al., 2004)
Dsp	Dead chimney	Southern Mariana Trough	12°55.15'N	143°36.96'E	Ipltc	271	73	26.9	(Kato et al., 2010)
Dsy	Dead chimney	Southern Mariana Trough	12°56.60'N	143°36.80'E	Iydc	144	43	29.9	(Kato et al., 2010)
Me	Mn crust	East Pacific Rise	9°30.36'N	104°13.48'W	3970-MO1A(MnO)	84	58	69.0	(Santelli et al., 2008)
Mt	Mn crust	Takuyo seamount	22°40.98'N	153°14.62'E	953Mn	86	65	75.6	(Nitahara et al., in press)
Re1	Basaltic rock	East Pacific Rise	9°28.48'N	104°15.0'W	3968-O8a(FeO2)	81	53	65.4	(Santelli et al., 2008)
Re2	Basaltic rock	East Pacific Rise	9°50.38'N	104°17.86'W	4059-B2(AlO)	74	40	54.1	(Santelli et al., 2008)
Re3	Basaltic rock	East Pacific Rise	9°49.41'N	104°16.32'W	3965-I2(FeO1)	52	46	88.5	(Santelli et al., 2008)
Re4	Basaltic rock	East Pacific Rise	9°49.42'N	104°16.32'W	3967-O2(FG)	79	43	54.4	(Santelli et al., 2008)
Re5	Basaltic rock	East Pacific Rise	9°50.78'N	104°17.58'W	9NBGBact	26	25	96.2	(Mason et al., 2007)
Rh1	Basaltic rock	Hawaii	18°58.31'N	155°53.42'W	PV550-X3(SP2)	91	46	50.5	(Santelli et al., 2008)
Rh2	Basaltic rock	Hawaii	18°58.31'N	155°53.42'W	PV550-X3(SP1)	64	44	68.8	(Santelli et al., 2008)
Rh3	Basaltic rock	Hawaii	18°52.17'N	155°14.53'W	PV547-X3(LSR)	71	25	35.2	(Santelli et al., 2008)
Rh4	Basaltic rock	Hawaii	18°55.02'N	155°15.53'W	PV549-X2(LPP)	246	148	60.2	(Santelli et al., 2008)
Rj	Basaltic rock	Juan de Fuca Ridge	46°41.95'N	130°55.94'W	JdFBGBact	21	21	100	(Mason et al., 2007)
Sdi	Sediment	Izu-Bonin Trench	30°55.00'N	141°49.00'E	BD5	18	9	50.0	(Li et al., 1999)
Sdt	Sediment	Takuyo seamount	22°40.98'N	153°14.62'E	953Sed	61	46	75.4	(Nitahara et al., in press)

Table 1. List of the clone libraries used in this study.

ecology. It is difficult to answer questions, like why the diversity of dead chimneys and seawater are lower than that of Mn crusts, basaltic rocks and sediments. Further data collection and experimental investigations are needed to answer.
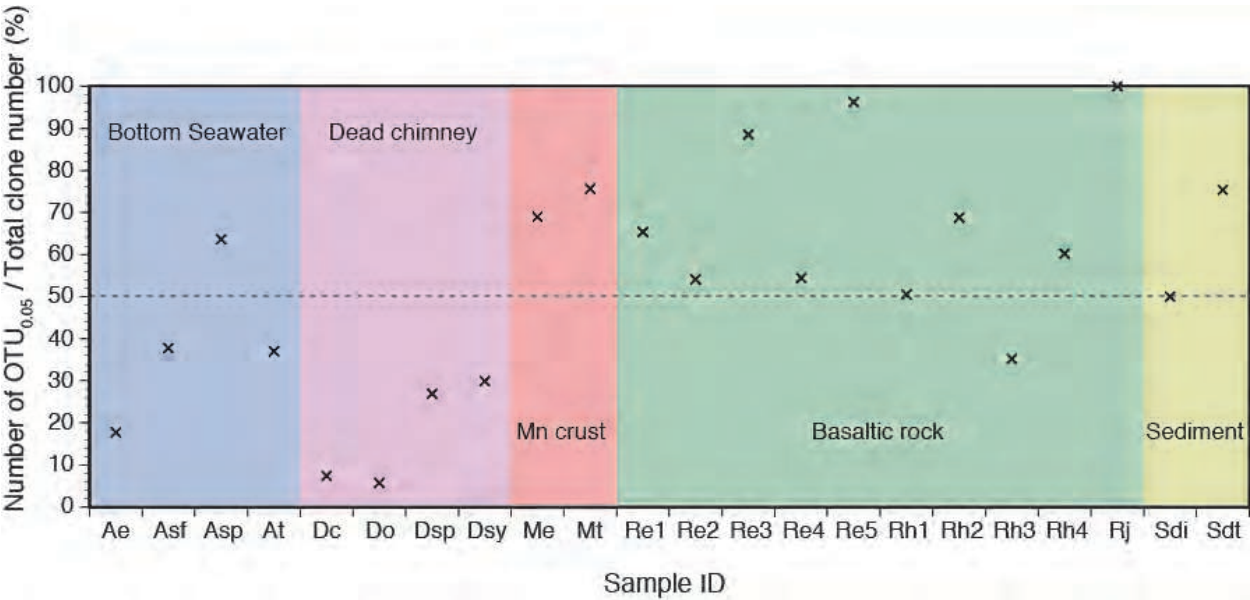

Fig. 3. The percentage of the number of $OTU_{0.05}$ in the total clone number for each sample.

Measurement and comparison of α-diversity is the first step in the study of biogeography. Investigating what kinds of habitats represent high or low diversity within a community is important for understanding the mechanism how microbial communities acquire the diversity. It should be noted that unification of the methodological process including DNA extraction and PCR primer sets in sequencing process (Wintzingerode et al., 1997), and alignment software, used region in the alignment dataset and distance calculation methods in phylogenetic and statistic processes (Schloss, 2010) is important to compare fairly the diversity among communities.

3.3 β-diversity

To compare the microbial communities for each sample based on the β-diversity measures, UniFrac was used here. The result of clustering analysis is shown in Figure 4A. In UniFrac analysis, the jackknifing method can be used to assess confidence in the nodes of the UPGMA tree. In the present case, all nodes, except the root, in the tree were not strongly supported by jackknifing (<50%). The PCoA results are shown in Figures 4B to E. Figure 4B is a three-dimensional image representing the first, second and third principal coordinate axes. Based on the results of clustering analysis and PCoA and taken each habitat-type into account, the samples were affiliated to six groups as shown in Figure 4. Group1 to Group3 represent the communities of basaltic rocks and Mn crusts and of each one of the sediment and chimney, respectively. Group4 and Group5 represent communities of dead chimneys and those of ambient bottom seawater, respectively. Group6 including As and Sdi is the most far from the other communities. Re5 was excluded from any groups due to the unambiguous behavior that may be caused by the small size of the clone libraries.

Using UniFrac, we can easily compare the samples and see the difference in community-level (Figure 4). The results indicated that seawater communities Group5 were clearly

distinguished from the other communities. However, the difference between Group4 (representing dead chimney communities) and Group5 were not shown by the two-dimensional image by the first and second axes (describing 10.81% and 7.79% of the variation, respectively; Figure 4C). On the other hand, these communities were separated along the third axis describing 6.93% of the variation. This means that the third most influential factor for the community similarity, not the first and second factors, is the factor separating the seawater and dead chimney communities. If certain environmental characteristics in habitats (e.g., pH and availability of energy sources such as iron, sulfide and ammonium as electron donors and oxygen, nitrate and sulfate as electron acceptors) and physiological characteristics of microbes in the communities (e.g., life styles of free-living and attachment) were correlated with the third axis, this traits would be the factor causing the difference between seawater and dead chimney communities. Likewise, details
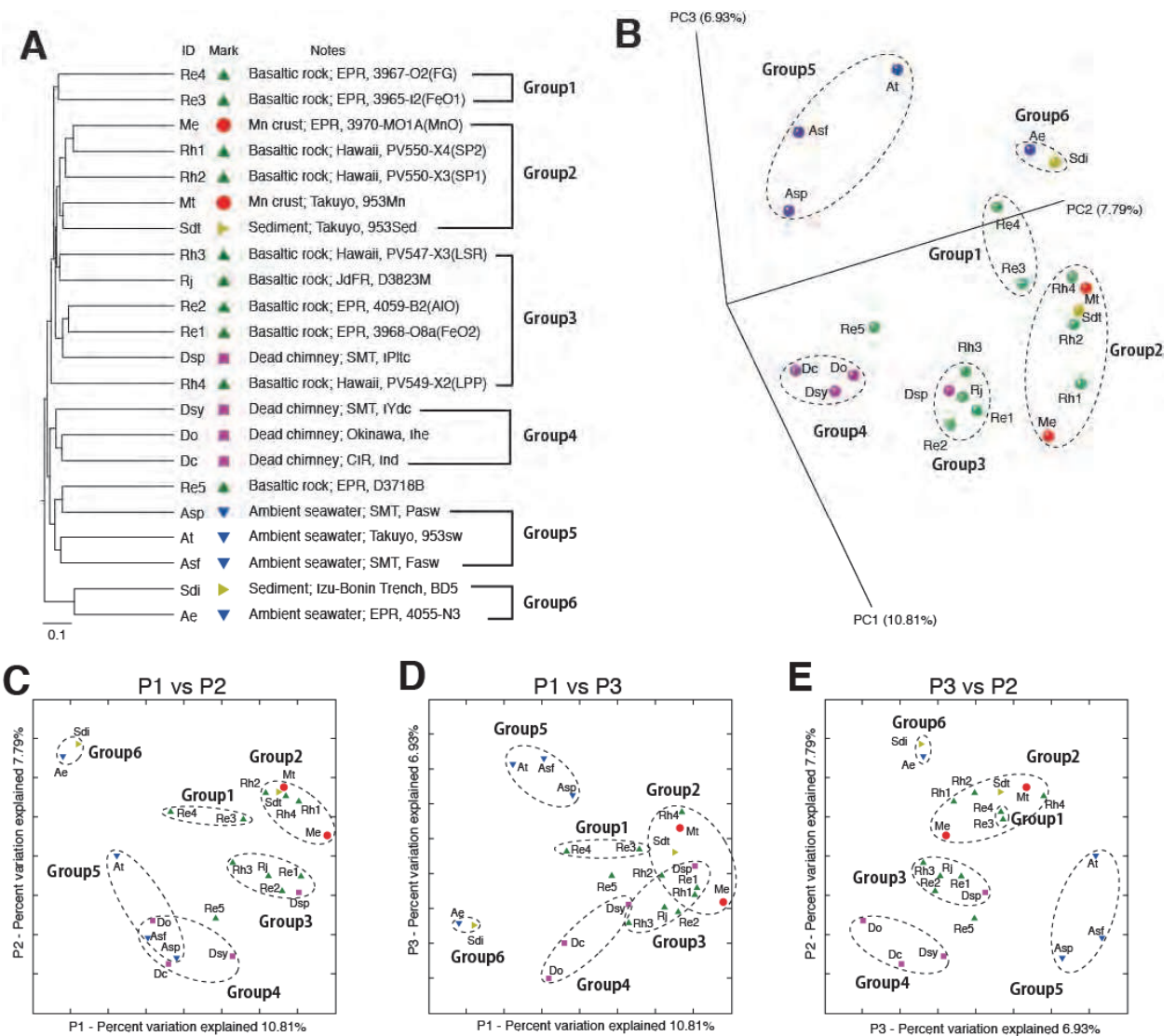

Fig. 4. The comparative results from (A) cluster analysis and (B-E) principal coordinate analysis using UniFrac. (B) The 3D image based on the first to third principal coordinates and (C-E) 2D images for the first vs. second, first vs. third, and second vs. third principal coordinates are shown, respectively.

of geographical and physicochemical characterization of the environments and physiological characterization of the communities in the habitats will help us to address the factors, for example, causing the difference among Group1 to Group4, causing the grouping of Group2 that contains Rh1, Mt and Sdt despite of the different habitat types (Mn crust, basaltic rock and sediment), and causing the difference between Group6 and the other groups.

Geographic distance may be one of the factors affecting the biogeography of the microbial communities. For example, the genetic distance between pairs of populations of *Synechococcus* or *Sulfolobus* in hot springs is related to the geographic distance despite the similar environmental characteristics of each habitat, which can be interpreted due to genetic drift caused by geographical isolation or to adaptation to the fluctuating environment in the past time (Papke et al., 2003; Whitaker et al., 2003). For the deep seafloor communities, the relationships between the community similarity and the geographic distance are shown in Figure 5. For all habitat types (i.e., basaltic rocks, Mn crusts, dead chimneys, sediments and bottom seawater), the similarity between the communities seems not to be related to the geographic distance (Figure 5A). For basaltic rocks and dead chimneys (Figure 5B and C), positive correlation between the community similarity and geographic distance is also not observed. In such solid habitats, environmental characteristics can be varied; for example, the gradient of oxygen concentration may occur due to chemical and biological consumption. The environmental varieties will affect on the microbial community diversity and composition in these habitats. This means that the community traits would be dramatically biased by its sampling position (e.g., interior or exterior parts of a sample). Hence, it is difficult to show the clear relationship between the community similarity and the geographic distance for solid habitats. For overlying bottom seawater (Figure 5D), our result implies that the community similarity is related to the geographic distance. Such correlation for marine microbial communities has been already reported (García-Martínez & Rodríguez-Valera, 2000). Further data sampling from various depths and locations in global oceans will provide a more clear view of oceanic biogeography.
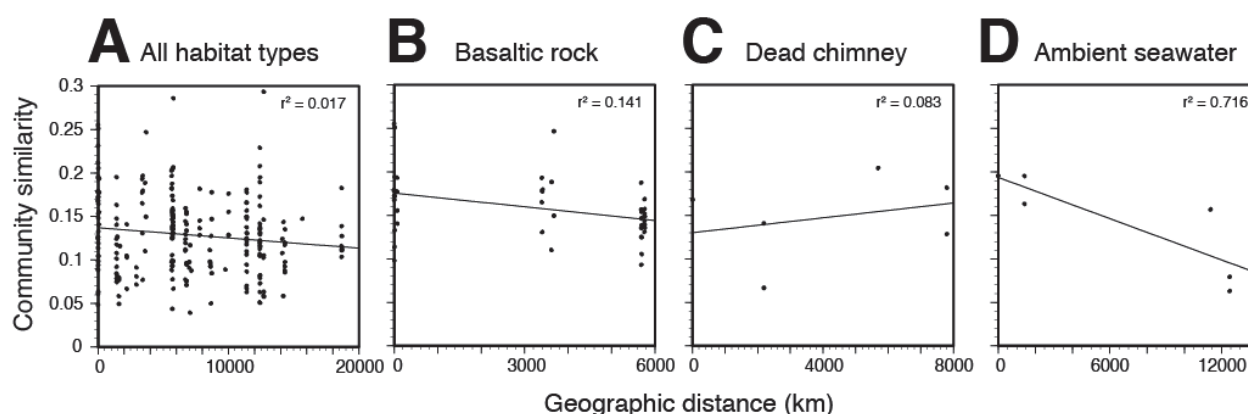

Fig. 5. The relationship between the community similarity and the geographic distance. The results of (A) all habitat types integrated, (B) basaltic rocks, (C) dead chimneys, and (D) ambient seawater are shown, respectively. The similarities between pairs of communities are calculated using UniFrac: the value 1 means that the two communities are the same. The geographic distance between pairs of habitats is calculated using spDistN1 in sp package of R software. Fitted line and coefficient regression value are shown in each figure.

To address the relationship between the community similarity and the phylogeny of the OTUs in each community, the OTUs in each community must be compared at nucleotide-sequence level, neither the band pattern of DGGE nor the fragment sizes of T-RFLP. As shown by the results of UniFrac (Figure 4), the community similarity on the deep seafloor has positively related to the habitat types. How many OTUs were shared among each community in genus- and family-levels (i.e., 0.05 and 0.20 cutoff) are shown in Venn diagrams (Figure 5) depicted by mothur.

Among the solid samples (i.e., basaltic rocks, Mn crusts, dead chimneys and sediments), the integrated community for the dead chimneys and that for the basaltic rocks contain many unique $OTUs_{0.05}$ (67.7% and 70.0%, respectively) in genus-level in contrast to that for sediments or Mn crusts (39.6% and 41.9%, respectively) (Figure 6A). In family-level, over 80% of $OTUs_{0.20}$ of each community were shared with others (Figure 6B). These unique $OTUs_{0.05}$ potentially contain indigenous members for each habitat; for example, the unique clusters for basaltic rocks, Ocean Crust Clades I to VII defined by Mason et al. (2007), and for sulfide chimney, Cluster A to C defined by Kato et al. (2010). Unfortunately, it is unclear whether and how these potential indigenous members play a role in the microbial ecosystem and elemental cycling because they are not phylogenetically close to known cultured species and their physiological characteristics are unknown (Mason et al., 2007; Kato et al., 2010). Further cultivation effort and characterization is important to link the phylogeny of the OTUs to their function and significance in the environments.

Over 50% of the $OTUs_{0.05}$ in the Mn crusts were shared with those in the basaltic rocks (Figure 6A). Furthermore, all of the $OTUs_{0.20}$ in the Mn crusts were shared with those in the basaltic rocks (Figure 6B). These results indicate that the bacterial members in the communities of the Mn crusts and basaltic rocks are phylogenetically close to each other, which is consistent with the UniFrac (divergence-based) result that the Mn crust communities clustered with some basaltic rock communities (Group2 in Figure 4). The phylogeny of the shared and unique OTUs can be confirmed by phylogenetic analysis (such as homology search against public databases and phylogenetic tree construction), although this is not shown in this chapter. Although the physiology of OTUs (e.g., metabolic function, growth rate and optimal growth temperature and pH) cannot be directly determined by their phylogeny, the physiological characteristics of OTUs may not be so different from those of certain cultured species that are closely related to the OTUs. The physiology of OTUs inferred from their phylogeny will provide basal information for constructing working hypothesis of the microbial ecosystem modeling and for preparing culture media targeting these uncultured members.

Such comparative analysis using nucleotide sequences are also used to check cross-contamination among each habitat. The shared OTUs between the bottom seawater community and others are shown in Figure 6C and D. In genus-level, approximately 6-8% of the total $OTUs_{0.05}$ of the Mn crusts, dead chimneys or basaltic rocks were shared with the seawater community (Figure 6C). In family-level, 83%-100% of $OTUs_{0.20}$ of each community were shared with others (Figure 6D), similar to the comparison among the solid samples (Figure 6B). Given that all of the $OTUs_{0.05}$ detected in the seawater are indigenous in the seawater, these shared $OTUs_{0.05}$ observed in the solid samples are potentially contaminants from the seawater community. However, it is also possible that these shared $OTUs_{0.05}$ are different from each other in higher-similarity level (e.g., 97% or 99% similarity). Microdiverse clusters at the level of >99% similarity have been reported for marine prokaryotes such as *Pelagibacter* (SAR11 cluster) (Acinas et al., 2004), Marine Group I

Crenarchaeota (Durbin & Teske, 2010), and for *Halomonas* and *Marinobacter* (Kaye et al., 2011). Hence, we need to be careful in concluding the shared OTUs between target and reference environments to the contaminants from the reference environment.
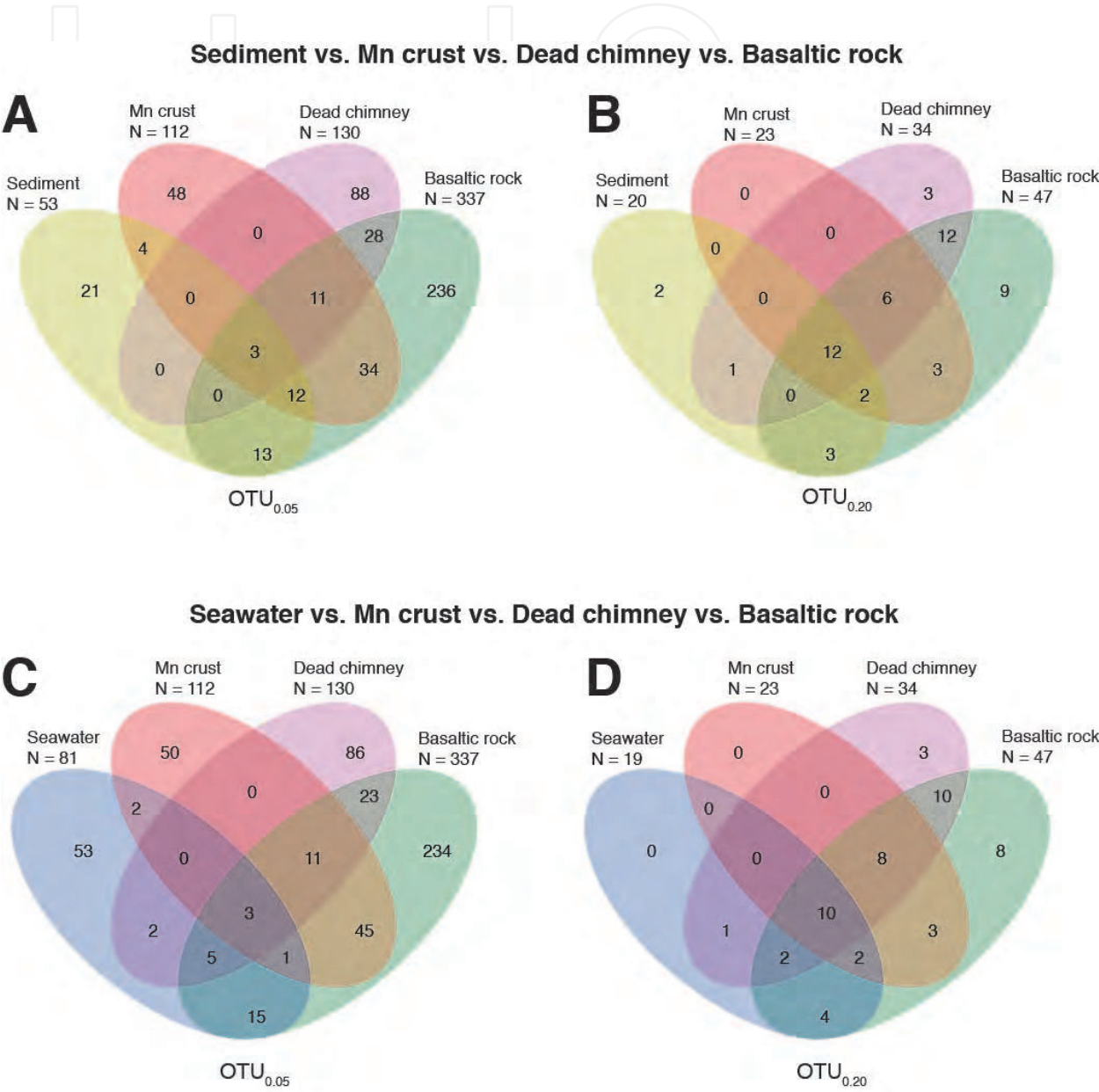

Fig. 6. Venn diagrams comparing the $OTU_{0.05}$ or $OTU_{0.20}$ memberships found in the sandy sediment, Mn crusts, dead chimneys, basaltic rocks and seawater samples. The comparative results among the sandy sediment, Mn crusts, dead chimneys and basaltic rocks are shown in (A) $OTU_{0.05}$ or (B) $OTU_{0.20}$ levels. The comparative results among the seawater, Mn crusts, dead chimneys and basaltic rocks are shown in (C) $OTU_{0.05}$ or (D) $OTU_{0.20}$ levels.

4. Concluding remarks

In this chapter, we introduce the recent bioinformatics tools for assessing the microbial diversity and biogeography. These useful tools allow us to analyze vast sequence data fast and correctly and to get the entire view of the biogeography of microbial communities in natural environments. We should use these tools for analysis of microbial biodiversity and biogeography effectively. Furthermore, both nucleotide sequencing technology and bioinformatics are developing steadily. Microbiologists, especially who study not only biogeography and biodiversity, but also evolution, ecology and biogeosciences, should always try not to overlook these advancing techniques and to apply to their studies. We applied the recent bioinformatics tools for actual data collected from deep seafloor environments. Our results provide insight into the microbial diversity and biogeography of the global deep seafloor in open oceans: for example, relationship between the community similarity and habitat types or geographic distance, commonality and difference among the communities in community- and OTU- levels. For providing persuasive explanation regarding the biogeography in the global deep seafloor, carefull collection of more molecular biological and environmental data from more seafloor habitats in various locations are needed.

5. References

Acinas SG, Klepac-Ceraj V, Hunt DE, Pharino C, Ceraj I, Distel DL & Polz MF (2004) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430: 551-554.

Amann RI, Ludwig W & Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* 59: 143-169.

Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ & Weightman AJ (2006) New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Appl Environ Microbiol* 72: 5734-5741.

Baas Becking LGM (1934) *Geobiologie of inleiding tot de milieukunde*. The Hague, the Netherlands: WP Van Stockum & Zoon.

Burggraf S, Larsen N, Woese CR & Stetter KO (1993) An intron within the 16S ribosomal RNA gene of the archaeon pyrobaculum aerophilum. *Proceedings of the National Academy of Sciences of the United States of America* 90: 2547-2550.

Cho J-C & Tiedje JM (2000) Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. *Appl Environ Microbiol* 66: 5448-5456.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM et al. (2009) The ribosomal database project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* 37: D141-D145.

Darling KF, Wade CM, Stewart IA, Kroon D, Dingle R & Brown AJL (2000) Molecular evidence for genetic mixing of arctic and antarctic subpolar populations of planktonic foraminifers. *Nature* 405: 43-47.

DeSantis TZ, Jr., Hugenholtz P, Keller K, Brodie EL, Larsen N, Piceno YM, Phan R & Andersen GL (2006) Nast: A multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucl Acids Res* 34: W394-399.

Duffy JE & Stachowicz JJ (2006) Why biodiversity is important to oceanography: Potential roles of genetic, species, and trophic diversity in pelagic ecosystem processes. *Marine Ecology Progress Series* 311: 179-189.

Durbin AM & Teske A (2010) Sediment-associated microdiversity within the Marine Group I Crenarchaeota. *Environmental Microbiology Reports* 2: 693-703.

Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792-1797.

Edwards KJ, Bach W & McCollom TM (2005) Geomicrobiology in oceanography: Microbe-mineral interactions at and below the seafloor. *Trends in Microbiology* 13: 449-456.

Excoffier, L. G. Laval, and S. Schneider (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.

Felsenstein J (1989) Phylip - phylogeny inference package (version 3.2). *Cladistics* 5: 164-166.

Fierer N & Jackson RB (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* 103: 626-631.

García-Martínez J & Rodríguez-Valera F (2000) Microdiversity of uncultured marine prokaryotes: The SAR11 cluster and the marine archaea of group I. *Molecular Ecology* 9: 935-948.

Gold T (1992) The deep, hot biosphere. *Proc Natl Acad Sci USA* 89: 6045-6049.

Guindon Sp, Dufayard J-Fo, Lefort V, Anisimova M, Hordijk W & Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of phyml 3.0. *Systematic Biology* 59: 307-321.

Hamady M, Lozupone C & Knight R (2009) Fast unifrac: Facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and phylochip data. *ISME J* 4: 17-27.

Hoagland P, Beaulieu S, Tivey MA, Eggert RG, German C, Glowka L & Lin J (2010) Deep-sea mining of seafloor massive sulfides. *Marine Policy* 34: 728-732.

Huber T, Faulkner G & Hugenholtz P (2004) Bellerophon: A program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* 20: 2317-2319.

Inagaki F, Sakihama Y, Inoue A, Kato C & Horikoshi K (2002) Molecular phylogenetic analyses of reverse-transcribed bacterial rRNA obtained from deep-sea cold seep sediments. *Environ Microbiol* 4: 277-286.

Itoh T, Suzuki K & Nakase T (1998) Occurrence of introns in the 16S rRNA genes of members of the genus thermoproteus. *Archives of Microbiology* 170: 155-161.

Itoh T, Nomura N & Sako Y (2003) Distribution of 16S rRNA introns among the family Thermoproteaceae and their evolutionary implications. *Extremophiles* 7: 229-233.

Kato S, Itoh T & Yamagishi A (2011) Archaeal diversity in a terrestrial acidic spring field revealed by a novel PCR primer targeting archaeal 16S rRNA genes. *FEMS Microbiology Letters* 319: 34-43.

Kato S, Kobayashi C, Kakegawa T & Yamagishi A (2009a) Microbial communities in iron-silica-rich microbial mats at deep-sea hydrothermal fields of the Southern Mariana Trough. *Environmental Microbiology* 11: 2094-2111.

Kato S, Takano Y, Kakegawa T, Oba H, Inoue K, Kobayashi C, Utsumi M, Marumo K, Kobayashi K, Ito Y et al. (2010) Biogeography and biodiversity in sulfide structures of active and inactive vents at deep-sea hydrothermal fields of the Southern Mariana Trough. *Appl Environ Microbiol* 76: 2968-2979.

Kato S, Yanagawa K, Sunamura M, Takano Y, Ishibashi J, Kakegawa T, Utsumi M, Yamanaka T, Toki T, Noguchi T et al. (2009b) Abundance of *Zetaproteobacteria* within crustal fluids in back-arc hydrothermal fields of the Southern Mariana Trough. *Environmental Microbiology* 11: 3210-3222.

Kaye JZ, Sylvan JB, Edwards KJ & Baross JA (2011) *Halomonas* and *Marinobacter* ecotypes from hydrothermal vent, subseafloor and deep-sea environments. *FEMS Microbiology Ecology* 75: 123-133.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.

Li L, Kato C & Horikoshi K (1999) Bacterial diversity in deep-sea sediments from different depths. *Biodiversity and Conservation* 8: 659-677.

Lozupone C & Knight R (2005) Unifrac: A new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71: 8228-8235.

Lozupone C, Hamady M & Knight R (2006) Unifrac - an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* 7: 371.

Lozupone C, Lladser ME, Knights D, Stombaugh J & Knight R (2010) Unifrac: An effective distance metric for microbial community comparison. *ISME J*.

Lozupone CA & Knight R (2008) Species divergence and the measurement of microbial diversity. *FEMS Microbiology Reviews* 32: 557-578.

Ludwig W, Strunk O, Klugbauer S, Klugbauer N, Weizenegger M, Neumaier J, Bachleitner M & Schleifer KH (1998) Bacterial phylogeny based on comparative sequence analysis. *Electrophoresis* 19: 554-568.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G et al. (2004) ARB: A software environment for sequence data. *Nucl Acids Res* 32: 1363-1371.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR et al. (2006) Microbial biogeography: Putting microorganisms on the map. *Nat Rev Micro* 4: 102-112.

Mason OU, Stingl U, Wilhelm LJ, Moeseneder MM, Di Meo-Savoie CA, Fisk MR & Giovannoni SJ (2007) The phylogeny of endolithic microbes associated with marine basalts. *Environmental Microbiology* 9: 2539-2550.

Nitahara S, Kato S, Urabe T, Usui A & Yamagishi A (2011) Molecular characterization of the microbial community in hydrogenetic ferromanganese crusts of the takuyo-daigo seamount, northwest Pacific. *FEMS Microbiol letters*. 321: 121-129 (2011)

Nocker A, Burr M & Camper A (2007) Genotypic microbial community profiling: A critical technical review. *Microbial Ecology* 54: 276-289.

Papke RT, Ramsing NB, Bateson MM & Ward DM (2003) Geographical isolation in hot spring cyanobacteria. *Environmental Microbiology* 5: 650-659.

Price MN, Dehal PS & Arkin AP (2010) FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5: e9490.

Prosser JI, Bohannan BJM, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, Green JL, Green LE, Killham K, Lennon JJ et al. (2007) The role of ecological theory in microbial ecology. *Nature Reviews Microbiology* 5: 384-392.

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J & Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* 35: 7188-7196.

Rona PA (2003) Resources of the sea floor. *Science* 299: 673-674.

Santelli CM, Orcutt BN, Banning E, Bach W, Moyer CL, Sogin ML, Staudigel H & Edwards KJ (2008) Abundance and diversity of microbial life in ocean crust. *Nature* 453: 653-656.

Schauer R, Bienhold C, Ramette A & Harder J (2009) Bacterial diversity and biogeography in deep-sea surface sediments of the south atlantic ocean. *ISME J* 4: 159-170.

Schloss PD (2010) The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS Comput Biol* 6: e1000844.

Schloss PD & Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* 71: 1501-1506.

Schloss PD & Handelsman J (2006) Introducing SONS, a tool for operational taxonomic unit-based comparisons of microbial community memberships and structures. *Appl Environ Microbiol* 72: 6773-6779.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75: 7537-7541.

Sheneman L, Evans J & Foster JA (2006) Clearcut: A fast implementation of relaxed neighbor joining. *Bioinformatics* 22: 2823-2824.

Smith WHF & Sandwell DT (1997) Global sea floor topography from satellite altimetry and ship depth soundings. *Science* 277: 1956-1962.

Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, Arrieta JM & Herndl GJ (2006) Microbial diversity in the deep sea and the underexplored "Rare biosphere". *PNAS*: 0605127103.

Suzuki Y, Inagaki F, Takai K, Nealson KH & Horikoshi K (2004) Microbial diversity in inactive chimney structures from deep-sea hydrothermal systems. *Microbial Ecology* 47: 186-196.

Takai K & Horikoshi K (1999) Genetic diversity of archaea in deep-sea hydrothermal vent environments. *Genetics* 152: 1285-1297.

R Development Core Team (2011) R: A language and environment for statistical computing. In. Vienna, Austria, http://www.R-project.org.

Webb CO, Ackerly DD & Kembel SW (2008) Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24: 2098-2100.

Whitaker RJ, Grogan DW & Taylor JW (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* 301: 976-978.

Whitman WB, Coleman DC & Wiebe WJ (1998) Prokaryotes: The unseen majority. *Proc Natl Acad Sci USA* 95: 6578-6583.

Wintzingerode F, Gobel UB & Stackebrandt E (1997) Determination of microbial diversity in environmental samples: Pitfalls of PCR-based rRNA analysis. *FEMS Microbiology Reviews* 21: 213-229.

Yannarell AC & Triplett EW (2005) Geographic and environmental sources of variation in lake bacterial community composition. *Appl Environ Microbiol* 71: 227-239.

Changing Diversity in Changing Environment

Edited by PhD. Oscar Grillo

As everybody knows, the dynamic interactions between biotic and abiotic factors, as well as the anthropic ones, considerably affect global climate changes and consequently biology, ecology and distribution of life forms of our planet. These important natural events affect all ecosystems, causing important changes on biodiversity. Systematic and phylogenetic studies, biogeographic distribution analysis and evaluations of diversity richness are focal topics of this book written by international experts, some even considering economical effects and future perspectives on the managing and conservation plans.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds