Isolated communities of *Epsilonproteobacteria* in hydrothermal vent fluids of the Mariana Arc seamounts

Julie A. Huber[*1], Holly V. Cantin[1], Susan M. Huse[1], David B. Mark Welch[1], Mitchell L. Sogin[1], and David A. Butterfield[2]

[1]Josephine Bay Paul Center, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543

[2]Joint Institute for the Study of Atmosphere and Oceans, University of Washington, PO Box 354925, Seattle, WA 98105

Running title: *Epsilonproteobacteria* in hydrothermal vents of the Mariana Arc

*Corresponding author. Phone: (508) 289-7291. Fax: (508) 457-4727. E-mail: *jhuber@mbl.edu*

ABSTRACT

Low-temperature hydrothermal vent fluids represent access points to diverse microbial communities living in oceanic crust. This study examined the distribution, relative abundance, and diversity of *Epsilonproteobacteria* in 14 low-temperature vent fluids from 5 volcanically active seamounts of the Mariana Arc using a 454 tag sequencing approach. Most vent fluids were enriched in cell concentrations compared to background seawater, and quantitative PCR results indicated all fluids were dominated by bacteria. Operational taxonomic unit (OTU)-based statistical tools applied to 454 data show that all vents from the northern end of the Marian Arc grouped together, to the exclusion of southern arc seamounts, which were as distinct from one another as they were from northern seamounts. Statistical analysis also showed a significant relationship between seamount and individual vent groupings, suggesting that community membership may be linked to geographical isolation and not geochemical parameters. However, while there may be large-scale geographic differences, distance is not the distinguishing factor in microbial community composition. At the local scale, most vents host a distinct population of *Epsilonprotoebacteria*, regardless of seamount location. This suggests there may be barriers to exchange and dispersal for these vent endemic microorganisms at hydrothermal seamounts of the Mariana Arc.

INTRODUCTION

Historically, biological studies of seamounts have focused on endemism and the hypothesis that some seamounts may host isolated macrobiological communities, similar to islands (Whitaker, 1998). While this theory remains controversial for macrofauna (Samadi*, et al.*, 2006, McClain, 2007, Stocks & Hart, 2007, Brewin*, et al.*, 2009), few studies have examined the issue of microbial endemism at seamounts. In the last 15 years, the microbiology of hydrothermally active seamounts has grown considerably (Emerson & Moyer, 2010), with the first reports of bacteria and archaea in microbial mats at Loihi Seamount near Hawaii (Moyer*, et al.*, 1994, Moyer*, et al.*, 1995) to diverse studies of different microbial habitats at Axial in the northeast Pacific (Huber*, et al.*, 2007, Opatkiewicz*, et al.*, 2009), Vailulu'u in the central Pacific (Staudigel*, et al.*, 2006), Suiyo in the northwest Pacific (Higashi*, et al.*, 2004), South Chamorro in the western Pacific (Mottl*, et al.*, 2003), and seamounts of the Mariana Arc in the western Pacific (Davis & Moyer, 2008). The chemistry of hydrothermally active seamounts can vary greatly depending on geological setting (forearc, backarc, volcanic arc, mid-ocean ridge, hotspot), making comparisons between different volcanoes challenging. Davis and Moyer (Davis & Moyer, 2008) presented the most comprehensive view of microbial diversity in hydrothermal seamounts to date based on their collection of microbial mats from Axial, Loihi, and volcanoes of the Mariana Arc and backarc. They found either *Epsilon-* or *Zeta-proteobacteria* (as assessed by T-RFLP) at most of the sites, and higher diversity on the arc systems compared to Axial and Loihi (mid-ocean ridge and hotspot). No biogeographical patterns were noted, but links between the microbial community structure and chemistry, depth, or temperature of the sample were suggested, demarcating three clusters of microbial communities. Often, multiple samples from within a seamount fell into more than one cluster,

highlighting the heterogeneity of the microbial mats and hydrothermal systems they sampled (Davis & Moyer, 2008). They concluded that arc and backarc hydrothermal systems may represent bacterial hotspots of diversity in the ocean.

Here, we focus on volcanoes of the Mariana Arc using venting fluids as an access point to the subseafloor microbial communities of seamounts (Schrenk*, et al.*, 2010). Our previous work in diffuse vents from mid-ocean ridges showed remarkable microbial diversity (based on 16S rDNA sequencing) and evidence that vents with different geochemical characteristics harbored distinct subseafloor microbial populations (Huber*, et al.*, 2006, Huber*, et al.*, 2007). This study uses a 454 tag sequencing approach to examine the distribution and diversity of the *Epsilonproteobacteria* in the subseafloor of hydrothermally active seamounts of the Mariana Arc. The *Epsilonproteobacteria* are one of the most dominant groups of microorganisms found in almost all hydrothermal vent biotopes, including microbial mats, sulfide chimneys, diffuse fluids, and associated with animals (Reysenbach*, et al.*, 2000, Corre*, et al.*, 2001, Huber*, et al.*, 2003, Higashi*, et al.*, 2004, Nakagawa*, et al.*, 2005, Takai*, et al.*, 2005, Campbell*, et al.*, 2006, Moussard*, et al.*, 2006, Huber*, et al.*, 2007). After initial detection by molecular methods, the first members of this ubiquitous group were isolated from hydrothermal vents in 2001 and since then have been cultured from vents around the world (Campbell*, et al.*, 2001, Takai*, et al.*, 2005). They display wide metabolic and phylogenetic diversity, but in general most are hydrogen- and sulfur-oxidizing chemolithoautotrophs, including both mesophiles and thermophiles (Campbell*, et al.*, 2006). All *Epsilonproteobacteria* studied to date from deep-sea vents fix carbon via the reductive or reverse tricarboxylic acid (rTCA) cycle (Campbell*, et al.*, 2001, Takai*, et al.*, 2005). Because they are not found in high abundances in oxygenated deep seawater and are very

abundant in hydrothermal vents, the *Epsilonproteobacteria* serve as an ideal indicator organism for examining vent-specific microbial community structure and distribution patterns.

MATERIALS AND METHODS

Field Sites and Sampling

In 2006, as part of the NOAA Ocean Explorer Ring of Fire program (Embley*, et al.*, 2004), we sampled 14 vents from 5 different seamounts along the Mariana Arc (Fig. 1, Table 1) using the Hydrothermal Fluid and Particle Sampler (HFPS) (Butterfield*, et al.*, 2004) on the ROV *Jason 2*. Fluids for DNA-based analyses passed through a Sterivex-GP (0.22-µm pore size) filter which was frozen at -80 ºC upon recovery of the vehicle. Whole unfiltered fluids for chemical analyses were also collected and analyzed according to Butterfield *et al*. 2004. Fluid samples were also preserved for cell enumeration using epifluorescence microscopy with DAPI as described previously (Huber *et al*., 2002), as well as cultivation of anaerobic thermophiles and hyperthermophiles.

Quantitative PCR

Quantitative polymerase chain reaction (qPCR) TaqMan assays previously described by Nadkarni *et al*. (2002) and Takai and Horikoshi (2000) were used to determine the relative abundance of bacterial and archaeal 16S rRNA genes in environmental vent fluids. Plasmid DNA was extracted from Axial Seamount low-temperature diffuse vent clone libraries, purified, and linearized using the WizardPlus SV Minipreps DNA Purification System (Promega). Standards were constructed by mixing equal amounts of four bacterial plasmids for quantification of bacterial 16S rRNA and two archaeal plasmids for quantification of archaeal 16S rRNA. A 1:10 dilution series of the plasmid mixtures beginning with an initial concentration of approximately 0.008 ng/µl was used to produce standard curves with R^2 values greater than

0.998 and efficiencies ranging from 86% to 99%. Each 20 µl reaction contained TaqMan Gene Expression Master Mix (Applied Biosystems), forward and reverse primers at optimized concentrations of either 9 µM (Bacteria) or 8 µM (Archaea), optimized probe concentrations of either 1.5 µM (Bacteria) or 2.0 µM (Archaea), DEPC-treated water, and 2 µl of DNA template. Reactions were performed in triplicate including no template controls on a StepOne Plus Real Time PCR System (Applied Biosystems). Cycles began with initial denaturation for 2 minutes at 50°C and 10 minutes at 96°C, followed by 40 cycles of 15 seconds at 95°C and 3 minutes at 59°C. StepOne Software Version 2.0 (Applied Biosystems) was used to analyze the results.

454 Tag Sequencing

Total genomic DNA was extracted according to Sogin *et al*. 2006. DNA was quantified using a Nanodrop ND-1000 spectrophotometer, and PCR amplicons of the bacterial V6 region generated ausing the pool of 5 forward primers and 4 reverse primers with a barcoding approach as described in Huber *et al*. 2007. All sequencing was carried out on the Roche GS20.

Sequence and Statistical Analysis

Pyrosequencing reads were passed through quality filters to reduce the error rate (Huse*, et al.*, 2007). Any read shorter than 50 nucleotides or containing one or more ambiguous nucleotides was discarded. The expected sample barcode and primer sequences were trimmed from the proximal and distal ends of the reads. Reads lacking a perfect match to the forward primer or a recognizable match to the reverse primer sequence (including rare variants) at either end, reads covering only a portion of the variable region, and reads that could not be unambiguously assigned to a sample were discarded. The resulting datasets contained 229,913 V6 tags. Sequences were assigned taxonomy using the Global Alignment for Sequence Taxonomy (GAST) process (Huse*, et al.*, 2008), which incorporates the RDP II taxonomy (Cole*,*

et al., 2007). 74,767 tag sequences mapped to the *Epsilonproteobacteria*. Sequences have been deposited in the VAMPS database (https://vamps.mbl.edu) and the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the accession number SRA012278.

A variety of analyses were then carried out using internal SQL database queries, SLP PWAL (Huse*, et al.*, 2010), mothur (Schloss*, et al.*, 2009), and Primer-E (Clarke and Gorley, 2006). All *Epsilonproteobacteria* sequences from each sample were pooled and OTUs calculated via SLP PWAL and mothur. To reduce the residual effects of sequencing error and noise, SLP PWAL uses pairwise alignments of the V6 tags based on the Needleman-Wunsch implementation of ESPRIT (Sun*, et al.*, 2009) to precluster the tags at the 2% difference level (corresponding to a single difference in the V6 region) using a modified single-linkage method. Each precluster is represented by its most abundant sequence in the final clustering step, which uses the ESPRIT pairwise alignment distances and the average-linkage method of mothur. The resulting matrix of OTUs shared between samples was normalized to total and transformed via square root and used to calculate a distance matrix with the Morisita-Horn measure (Horn, 1966). Distance matrices were imported into Primer-E for a variety of analyses, including hierarchical cluster analysis, principal component analysis, ANOSIM, SIMPROF and non-metric multi-dimensional scaling (MDS) analysis. Analyses were also repeated with all singletons removed. The entire chemical dataset, including temperature, pH, gas (H_2S, H_2, CH_4), Na, Li, Fe, Mn, Si was also included for PCA and Spearman correlation analyses.

RESULTS

Site Descriptions

The Mariana Arc is located in the western Pacific from about 14 º N to 24 º N (Bloomer, *et al.*, 1989) (Fig. 1). We analyzed vent fluids from 5 hydrothermally active seamounts along the Arc: NW Rota-1 and Forecast are located in the Southern Seamount Province, while Daikoku, NW Eifuku, and Nikko lie in the Northern Seamount Province (Stern, *et al.*, 1988, Embley, *et al.*, 2004) (Fig. 1, Table 1). All of the seamounts occur along the active Mariana Arc front, with the exception of Forecast, which may be more closely affiliated with the back-arc spreading axis, although this remains unknown (Embley, *et al.*, 2004). At all five sites, we sampled fluids that were venting directly out of the seafloor as opposed to those venting from sulfide structures. There are a number of excellent geological descriptions and plume water surveys of the volcanoes sampled elsewhere (Embley, *et al.*, 2004, Embley, *et al.*, 2007, Baker, *et al.*, 2008, Resing, *et al.*, 2009), therefore here we will only summarize the specific vent sites sampled.

Moving south to north, the first seamount sampled was Forecast, previously discovered by the Japanese in 1993 (Gamo & Cruise, 1993, Ishibashi & Urabe, 1995). At this site, we found chimneys vigorously venting clear fluids up to 195 ºC, one of the highest temperature vent systems known in the Mariana Arc. The two samples used in our analysis were from low temperature vents located 80 meters apart on the summit. Snail Scrum (FS431) was taken in a group of snails about 1 meter away from a high temperature structure, and Homer Vent (FS432) was taken in a rock crack, also a few meters away from a different high temperature chimney. Both of these vents had relatively high pH values (Table 1) and low gas (H_2, H_2S, and CH_4) concentrations (Marvin Lilley, unpublished data). The second site sampled was NW Rota-1, an actively erupting volcano with abundant diffuse venting (Embley, *et al.*, 2006, Chadwick, *et al.*, 2008). We sampled the eruptive Brimstone pit (FS445) during a quiet period, as well as 4 diffuse vents (FS446-449) trending east and downslope from the eruptive area, many of which

had abundant white microbial mat, as well as shrimp present. The lowest pH values were found both at the eruptive pit (pH 1.64) and at nearby diffuse vents Sandy Saddle and Iceberg (pH 2.62 and 3.75, respectively). Concentrations of hydrogen and methane were the highest seen on the arc (Marvin Lilley, unpublished data). In the Northern Seamount Province, we sampled Daikoku, an unusual site with ponds of molten sulfur and thick sulfur crusts, as well as abundant macrofauna, including tube worms, snails, and flatfish (Embley*, et al.*, 2007). Alka Seltzer vent (FS475) was located near the molten sulfur cauldron and gas bubbles were visible during sampling; NE Pit (FS473) was at the summit, where a plume of white smoke was sampled (FS473). Nikko, the northern-most site, was similar in venting style and macrofauna to Daikoku and hosted pillars of native sulfur, small "pots" of molten sulfur bubbling out of the seafloor, and extensive communities of tubeworms and other fauna (Embley*, et al.*, 2007). Site North vent (FS479) was the highest temperature vent sampled, close to 100 ºC, and was a plume of white smoke; Tubeworm Hangover (FS480), an area of dense microbial mat and animals; and Top Vent (FS481), a clear venting fluid with abundant microbial mat and shrimp. Both Nikko and Daikoku fluids had pH values around 5 (although Alka Seltzer had a pH of 3.36) with relatively low gas and iron concentrations (Marvin Lilley, unpublished data). The final site was NW Eifuku, the deepest of the sites, with venting fluids rich in both liquid and gas CO_2, as well as large mussel beds and dense microbial mats (Lupton*, et al.*, 2006, Embley*, et al.*, 2007, Tunnicliffe*, et al.*, 2009). Both sites sampled here, Sulfur Flow (FS467) and Cliff House (FS468) were taken near areas where liquid CO_2 droplets were venting with high concentrations of CO_2 and H_2S as well as minor CH_4 and H_2 (Lupton*, et al.*, 2006). The gas data (CO_2, H_2S, H_2, CH_4) are presented in Lupton et al. (2006) and further discussed in Lupton et al. (2008).

Microbial Communities in Fluids: Cell counts, qPCR, and 454 Sequencing

All fluids sampled were enriched in cell densities compared to background seawater at the same depth, with the exception of North Vent (FS479) on Nikko and Iceberg (FS446) on NW Rota-1 (Table 1). Quantitative PCR results indicate all fluids were dominated by bacteria, with archaea compromising less than 3% of the total community, although Brimstone Pit (FS445) on NW Rota-1 had almost 8% archaea (Table 2). Cultivation of thermophiles and hyperthermophiles in anaerobic media was attempted on most samples; out of approximately 250 inoculations, 16 tubes were verified for positive growth via microscopy. These cultures are being examined as part of a separate study. Total DNA was extractable from all fluids, and amplification of the bacterial V6 region successful. Overall, 229,913 bacterial sequences were obtained from the 14 samples, 74,767 of which mapped to the *Epsilonproteobacteria* after the GAST process (Huse*, et al.*, 2008) (Table 2). The relative abundance of *Epsilonproteobacteria* in the total bacterial sequences per sample (Table 2) ranged from 3.7% in NE Pit (FS473) from Daikoku to 87.4% in Tubeworm Hangover (FS480) from Nikko. No correlation between relative abundance and location or chemical parameter was found. All subsequent analyses were carried out on this dataset of 74,767 sequences.

Taxonomy and Total Diversity of Epsilonproteobacteria

The taxonomic breakdown at the genera level for each sample is shown in Figure 2. Dominant genera included *Thioreductor*, *Lebetimonas*, *Caminibacter*, *Sulfurovum*, *Sulfurimonas* and *Arcobacter*. Sequences that were found less than 1000 times across all samples were binned into "Other" and included *Nautilia*, *Sulfuricurvum*, *Hydrogenimons*, *Nitratifractor*, and *Sulfurospirillum*. Because sites were unequally sampled, rarefaction was used to compare sampling effort across all sites at the 3% difference level (Fig. 3). Results indicate that 8 of the 14 sites cluster together, whereas the two vents from Forecast have the highest diversity (FS431

and FS432) and two of the vents from NW Rota-1 along with Alka Seltzer from Daikoku have the lowest diversity (FS445, FS446, and FS475). It is only in these latter three sites that sampling of *Epsilonprotoebacteria* appears complete. Vents from Forecast had the highest pH and samples FS445, FS446, and FS475 had among the lowest with respect to pH values (Table 1).

Patterns of Epsilonproteobacteria Distribution

All sequences were pooled together to generate OTUs, resulting in 739 OTUs at the 3% difference level. Only 5 OTUs occurred in all 14 datasets, compromising less than 0.7% of the total OTUs but over 10% of the total sequences. Approximately 40% of the OTUs were found in 2 to 13 samples, compromising 88% of sequences. Over half of the OTUs (59%, representing about 0.6% of the total sequences) were only found in one sample. Similarity profile (SIMPROF) test of the data indicated that there was significant community structure differences among the individual samples ($\pi = 6.23$, p value <0.001). At the 3% difference level, MDS and cluster analysis found the pattern seen in Figure 4. The two vents from Forecast (FS431 and FS432) group together, the five vents from NW Rota-1 group together (FS445-FS449), and the vents from Nikko, Daikoku, and Eifuku group together. One-way ANOSIM analysis of similarity test using the distance matrix and seamount, resulted in a global R value of 0.667 ($p<0.0002$), indicating that the groupings by seamount are significant. Pairwise R values indicate that the major differences are found between the two southern seamounts (Forecast and NW Rota-1) and the southern seamounts as compared to all of the northern seamounts (NW Eifuku, Daikoku, and Nikko, Table 3). Nikko is the only seamount that has significant overlap with the other two northern seamounts, with North vent and Tubeworm Hangover (FS479 and FS480) clustering with Daikoku and Top vent (FS481) clustering with NW Eifuku (Fig. 4).

Sequences from each seamount were pooled and analysis repeated with grouping by seamount rather than individual vent site. A UPGMA tree calculated from Morisita-Horn distances between datasets (Fig. 5) supports the finding that Forecast and NW Rota-1 are distinct from one another and from the northern seamounts. No significant groupings or correlations were found based on chemical parameters, including temperature, pH, gas (H_2S, H_2, CH_4), Na, Li, Fe, Mn, and Si (data not shown).

Considerable concern has been raised recently over pyrosequencing errors and their contribution to diversity estimates (Quince*, et al.*, 2009, Reeder & Knight, 2009, Kunin*, et al.*, 2010). In order to account for potential errors, all samples were run on the same sequencing platform (GS20) within a few weeks of one another, and all samples were treated identically from DNA extraction all the way through to sequence analysis. We used stringent quality control measures based on our knowledge of known GS20 sequencing errors (Huse*, et al.*, 2007). In addition, here we report the most conservative methodology to date for reporting OTUs via the SLP PWAL pipeline (Huse*, et al.*, 2010). The SLP PWAL approach is thought to more accurately predict expected OTUs than multiple alignment and complete-linkage clustering methods and reduces the number of OTUs almost by half compared to the multiple alignment complete-linkage clustering methods (Huse*, et al.*, 2010). We also repeated all analysis with singletons removed to insure that potentially erroneous singletons were not skewing community patterns.

DISCUSSION

In this study, we used 454 tag sequencing of the V6 bacterial region to determine the diversity and distribution of *Epsilonproteobacteria* in rock-hosted diffuse vent fluids from 5 hydrothermally active seamounts along the Mariana Arc. While universal bacterial primers were

used and a large taxonomic diversity of bacteria was detected, we chose to only analyze a subset of the sequences that were classified to *Epsilonproteobacteria*. Diffuse vents represent a mixture of seawater and vent fluids that mix in the crust before venting on the seafloor (Butterfield*, et al.*, 2004). We sampled vent fluids from a large range of depths and locations (~400-1600m), and we did not want background seawater organisms from different regions and depths complicating interpretation of vent-specific patterns. *Epsilonproteobacteria* are rarely detected in non-plume, non-vent background oxygenated seawater and are considered endemic to sulfidic environments like deep-sea hydrothermal vents (Campbell*, et al.*, 2006).

Taxonomically, over 90% of the sequences that mapped to *Epsilonproteobacteria* were categorized down to the genus level, leaving only 9% of the sequences unclassified beyond *Epsilonproteobacteria*, although many were classified down to order and family. Members of *Thioreductor* spp. were found the most frequently. There is only one named species within this genus, *Thioreductor micantisoli*, and it was isolated from hydrothermal sediments in the Mid-Okinawa trough (Nakagawa*, et al.*, 2005). *T. micantisoli* is a strict autotrophic mesophile that oxidizes hydrogen while reducing sulfur, and sequences related to *Thioreductor* spp. have been retrieved from hydrothermal habitats throughout the worlds oceans, including the Mid-Atlantic Ridge associated with venting fluids (Reysenbach*, et al.*, 2000), the East Pacific Rise in microbial mats (Longnecker & Reysenbach, 2001), and the Indian Ocean associated with gastropods (Goffredi*, et al.*, 2004) The phylogenetic position of this genus remains undetermined and the physiological diversity completely unexplored. *Sulfurimonas* spp. and *Sulfurovum* spp. were also detected at high relative abundances. These are among the most frequently detected *Epsilonproteobacteria* in hydrothermal vent habitats, and both are mesophilic microaerobic autotrophs that oxidize elemental sulfur (Campbell*, et al.*, 2006) These

three genera—*Thioreductor*, *Sulfurovum*, and *Sulfurimonas*—were found at every site, although some at very low relative abundances. The remaining dominant genera, *Caminibacter* and *Lebetimonas*, both belong to the family *Nautiliaceae* and were not found at every site. This family has a number of cultured isolates, all of which are thermophilic autotrophs (or mixotrophs) that use hydrogen as an electron donor and both sulfur and nitrate as electron acceptors. They have been isolated from the East Pacific Rise, Mid-Atlantic Ridge, and the Mariana Arc, where the first acidophilic member was isolated, *Lebetimonas acidiphlia* (Takai*, et al.*, 2005). The largest group of unclassified sequences was found in sample FS446, where two tag sequences were found close to 3500 times and were only resolved to the family *Nautiliaceaea*. The best match revealed by the GAST process is a sequence from Davis and Moyer (2008) from microbial mats from the exact same vent, Iceberg on NW Rota-1. Their sequencing of full length 16S rDNA amplicons indicate this sequence falls into an unresolved group within the *Nautiliaceaea* (Accession No. EU574679, IBB_14). The taxonomic breadth found in the vent fluids also included rare members that occurred less than 1000 times total across the datasets, such as *Nautilia*, *Sulfuricurvum*, *Hydrogenimons*, *Nitratifractor*, and *Sulfurospirillum*.

Because samples were unequally sampled, we used rarefaction analysis rather than absolute OTU numbers or non-parametric estimators to compare diversity among samples. As illustrated in Figure 3, we found that the *Epsilonproteobacteria* in three of the samples (FS445, FS446, and FS475) appear to be the least diverse as indicated by the flattening of the rarefaction curve. Sample FS445 from NW Rota-1 has the lowest pH (1.64) of all vents sampled in this study, and samples FS446 and FS475 had low pHs as well (3.75 and 3.36, respectively). The other vent with a low pH (FS447, pH 2.62) did not display the same flattening trend. The two

vents with the highest community diversity as measured by rarefaction, FS431 and FS432, also had the highest pH values in this study (6.56 and 6.23, respectively). While the trend of pH and total diversity is not consistent across the entire dataset, there is some evidence to suggest that *Epsilonproteobacteria* diversity may be constrained by pH, similar to what is seen in soils (Lauber*, et al.*, 2009). However, no statistically significant relationship between pH and community composition was found. We used nMDS, cluster, PCA, correlation, SIMPROF and other statistical tools to examine the distribution patterns of *Epislonprotoebacteria* and how those patterns may be linked to environmental parameters, including 15 chemical species. Statistical tests (SIMPROF) indicate that there is evidence of significant group structure beyond what one would expect randomly, and the only significant relationship found was between seamount and individual vent groupings (ANOSIM, Table 3). Forecast and NW Rota-1 each formed distinct clusters, with very little to no overlap with each other or with vents found in the northern province of the arc (Fig. 4, Table 3). However, the three sites in the northern province group together, with subgroupings within. This suggests that community membership may be linked to geographical isolation, but it is important to note that at finer resolution and higher similarity values, many of the individual vents appear isolated. While there may be large-scale geographic differences between the northern and southern seamounts, at the local scale each vent is practically an island, with little community overlap with other vents, regardless of seamount location.

The concept of geographical isolation, rather than environmental parameters such as geochemistry, controlling microbial community structure has frequently been observed in microorganisms from terrestrial hot springs (Papke*, et al.*, 2003, Whitaker*, et al.*, 2003, Takacs-Vesbach*, et al.*, 2008, Reno*, et al.*, 2009), as well as soils and lakes (see (Papke & Ward, 2004,

Whitaker, 2006, Green*, et al.*, 2008) for reviews). In the marine hydrothermal environment, there is also some evidence of geographic isolation of microbial communities. For example, recent work at Axial Seamount found that individual vents within the caldera maintain a distinct composition over time (Opatkiewicz*, et al.*, 2009), and no correlation between total bacterial or archaeal community structure and geochemical parameters was found. However, the authors did find trends in *Epsilonproteobacteria* community composition that were linked to changes in hydrogen, sulfur, and iron chemistry. Similarly, Huber *et al*. (2003) found that *Epsilonproteobacteria* diversity increased over time in vent fluids from Axial Seamount following the 1998 eruption, and this increase followed changes in the mixing of seawater and temperature of the vent. Other qualitative relationships between chemistry and microbial community structure have been seen at deep-sea hydrothermal vents. Nakagawa *et al*. (2005) found that there were more chemolithoautotrophs, such as hydrogenotrophic methanogens, in gas-enriched fluids as compared to gas-depleted or normal vent fluids. Again working at Axial, Huber *et al*. (2006), found a correlation between the populations of Thermococcales (as defined by the ITS region) and vent chemistry. With the exception of Opatkiewicz *et al*. (2009), none of these studies used statistical tools to analyze relationships between chemistry and microbiology, and most did a very shallow level of sequencing in comparison to this study. The fact that no pattern was seen in our study may be due to a number of factors. It is possible that the bulk measurements carried out here for both chemistry and microbiology may obscure any meaningful relationships, and it is also well known that vent fluid chemistry changes over time. In addition, the level of mixing in arc seamounts is difficult to constrain because of the lack of zero-end member magnesium fluids (or other conservative tracers) and the extremely high volume of water rapidly moving through the rock. Another important factor is that we analyzed

only the DNA fraction of the community, meaning we measured all organisms present, regardless of activity. Patterns in community composition of active *Epsilonproteobacteria*, as studied by RNA, may reveal more significant relationships with environmental parameters. In addition, the V6 region of the ribosomal RNA may be too short or lack enough variability to detect fine-scale patterns. Using other portions of the genome, such as the intergenic spacer region or genes actively used for metabolism, may help resolve patterns further. Finally, other microorganisms may exhibit different patterns in their distribution and diversity, such as the *Zetaproteobacteria* (Rassa*, et al.*, 2009, Emerson & Moyer, 2010). However, our analysis shows that the biogeographic pattern of *Epsilonproteobacteria* at these seamounts is statistically significant.

Along the Mariana Arc, it is possible that the central province is a geographical barrier to distribution and exchange between the northern and southern provinces. The central province is mostly subaerial, with little submarine volcanism found (Embley*, et al.*, 2007). However, the two seamounts in the southern province (Forecast and NW Rota-1) were as different from one another as they were from seamounts to the north. Similarly, NW Eifuku and Daikoku have different community compositions, despite the fact that they are located closer together than any of the other seamounts. Clearly, distance is not the distinguishing factor in microbial community composition, and chemically, no statistically significant relationships were seen. While the chemistry of most of these arc volcanoes is dominated by magmatic gas and sulfur, there is still considerable chemical diversity at the vents, with wide ranges of pH, gas concentrations, and temperature, even within a single seamount. Therefore, it may not be surprising that each individual vent hosts its own distinct and somewhat isolated population of *Epsilonproteobacteria*. Telford *et al*. (2006) hypothesized that endemic organisms in rare or

isolated habitats have a lower probability of successful dispersal, thus allowing for niche specialization and genetic divergence. Hydrothermal vents are certainly isolated and often ephemeral habitats in the deep sea. At the level of resolution provided by 16S rDNA 454 tag sequencing, it appears that the pattern seen here is driven mainly by geographical isolation. Our data suggests that there may be barriers to exchange and dispersal for *Epsilonproteobacteria* at hydrothermal seamounts of the Mariana Arc. Further studies, including characterization of isolates, manipulative physiological experiments, and genomic characterizations, will help elucidate how to interpret the patterns seen in-situ. With the increasing amounts of sequencing data, we are now poised to carry out a much larger census of vent and seamount microorganisms to address the important issue of whether or not microbial endemism exists in benthic marine microbial communities.

ACKNOWLEDGEMENTS

We thank the NOAA Ocean Explorer Ring of Fire Program, the *ROV* Jason 2, and S. Murdock for field support and sample collection. This work was supported by a National Research Council Research Associateship Award and L'Oréal USA Fellowship (J.A.H.), NASA Astrobiology Institute Cooperative Agreement NNA04CC04A (M.L.S.), the Alfred P. Sloan Foundation's ICoMM field project, and the W. M. Keck Foundation. This publication is [partially] funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement No. NA17RJ1232, Contribution #1814. This is NOAA Pacific Marine Environmental Laboratory Contribution #3542.

REFERENCES

[1] Baker ET, Embley RW, Walker SL*, et al.* (2008) Hydrothermal activity and volcano distribution along the Mariana arc. *Journal of Geophysical Research* **113**: 1-16.

[2] Bloomer SH, Stern RJ & Smoot NC (1989) Physical volcanology of the submarine Mariana and Volcano Arcs. *Bulletin of Volcanology* **51**: 210-224.

[3] Brewin PE, Stocks KI, Haidvogel DB, Condit C & Gupta A (2009) Effects of oceanographic retention on decapod and gastropod community diversity on seamounts. *Marine Ecology Progress Series* **383**: 225-237.

[4] Butterfield DA, Lilley MD, Huber JA, Roe KK, Embley RW, Baross JA & Massoth GJ (2004) Mixing, reaction, and microbial activity in the sub-seafloor revealed by temporal and spatial variation in diffuse flow vents at Axial Volcano. *The subseafloor biosphere at mid-ocean ridges,* Vol. Geophysical Monograph 144 (Wilcock WSD, DeLong EF, Kelley DS, Baross JA & Cary SC, ed.^eds.), p.^pp. 269-289. American Geophysical Union, Washington, D.C.

[5] Campbell B, Jeanthon C, Kostka JE, Luther GWI & Cary SC (2001) Growth and phylogenetic properties of novel bacteria belonging to the epsilon subdivision of the *Proteobacteria* enriched from *Alvinella pompejana* and deep-sea hydrothermal vents. *Applied and Environmental Microbiology* **67**: 4566-4572.

[6] Campbell BJ, Engel AS, Porter ML & Takai K (2006) The versatile ε-proteobacteria:key players in sulphidic habitats. *Nature Reviews Microbiology* **4**: 458-468.

[7] Chadwick WW, Jr., Cashman KV, Embley RW*, et al.* (2008) Direct video and hydrophone observations of submarine explosive eruptions at NW Rota-1 volcano, Mariana arc. *Journal of Geophysical Research* **113**: 1-23.

[8] Cole JR, Chai B, Farris RJ*, et al.* (2007) The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Research* **35**: D169-D172.

[9] Corre E, Reysenbach A-L & Prieur D (2001) ε-proteobacterial diversity from a deep-sea hydrothermal vent on the Mid-Atlantic Ridge. *FEMS Microbiology Letters* **205**: 329-335.

[10] Davis RE & Moyer C (2008) Extreme spatial and temporal variability of hydrothermal microbial mat communities along the Mariana Island Arc and southern Mariana back-arc system. *Journal of Geophysical Research* **113**: 1-17.

[11] Embley RW, Baker ET, Chadwick WWJ, Lupton JE, Resing JA, Massoth GJ & Nakamura K (2004) Explorations of Mariana Arc volcanoes reveal new hydrothermal systems. *Eos Transactions, American Geophysical Union* **85**: 37-39.

[12] Embley RW, Baker ET, Butterfield DA*, et al.* (2007) Exploring the submarine ring of fire: Mariana Arc- Western Pacific. *Oceanography* **20**: 68-79.

[13] Embley RW, Chadwick WW, Baker ET*, et al.* (2006) Long-term eruptive activity at a submarine arc volcano. *Nature* **441**: 494.

[14] Emerson D & Moyer CL (2010) Microbiology of seamounts: common patterns observed in community structure. *Oceanography* **23**: 148-163.

[15] Gamo T & Cruise SSPotY (1993) Revisits to the mid-Mariana Trough hydrothermal site and discovery of new venting in the southern Mariana region by the Japanese submersible Shinkai 6500. *InterRidge News* **2**: 11-14.

[16] Goffredi SK, Warén A, Orphan VJ, Van Dover CL & Vrijenhoek RC (2004) Novel forms of structural integration between microbes and a hydrothermal vent gastropod from the Indian Ocean. *Applied and Environmental Microbiology* **70**: 3082-3090.

[17] Green JL, Bohannan BJM & Whitaker RJ (2008) Microbial biogeography: from taxonomy to traits. *Science* **320**: 1039-1043.

[18] Higashi Y, Sunamura M, Kitamura K*, et al.* (2004) Microbial diversity in hydrothermal surface to subsurface environments of Suiyo Seamount, Izu-Bonin Arc, using a catheter-type in situ growth chamber. *FEMS Microbiology Ecology* **47**: 327-336.

[19] Horn HS (1966) Measurement of 'overlap' in comparative ecological studies. *American Naturalist* **100**: 419-424.

[20] Huber JA, Butterfield DA & Baross JA (2003) Bacterial diversity in a subseafloor habitat following a deep-sea volcanic eruption. *FEMS Microbiology Ecology* **43**: 393-409.

[21] Huber JA, Butterfield DA & Baross JA (2006) Diversity and distribution of subseafloor Thermococcales populations in diffuse hydrothermal vents at an active deep-sea volcano in the northeast Pacific Ocean. *Journal of Geophysical Research* **111**: 1-13.

[22] Huber JA, Mark Welch DB, Morrison HG, Huse SM, Neal PR, Butterfield DA & Sogin ML (2007) Microbial population structures in the deep marine biosphere. *Science* **318**: 97-100.

[23] Huse SM, Welch DM, Morrison HG & Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* **10.1111/j.1462-2920.2010.02193.x**.

[24] Huse SM, Huber JA, Morrison HG, Sogin ML & Mark Welch DB (2007) Accuracy and quality of massively-parallel DNA pyrosequencing. *Genome Biology* **8**: R143.

[25] Huse SM, Dethlefsen L, Huber JA, Welch DM, Relman DA & Sogin ML (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics* **4**: e1000255.

[26] Ishibashi J & Urabe T (1995) Hydrothermal activity related to arc-backarc magmatism in the western Pacific. *Backarc Basins: Tectonics and Magmatism,*(Taylor B, ed.^eds.), p.^pp. 451-495. Plenum Press, New York.

[27] Kunin V, Engelbrektson A, Ochman H & Hugenholtz P (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology* **12**: 118-123.

[28] Lauber CL, Hamady M, Knight R & Fierer N (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and Environmental Microbiology* **75**: 5111-5120.

[29] Longnecker K & Reysenbach A-L (2001) Expansion of the geographic distribution of a novel lineage of ε-proteobacteria to a hydrothermal vent site on the Southern East Pacific Rise. *FEMS Microbiology Ecology* **35**: 287-293.

[30] Lupton J, Butterfield DA, Lilley M*, et al.* (2006) Submarine venting of liquid carbon dioxide on a Mariana Arc volcano. *Geochemistry Geophysics Geosystems* **7**: 1-20.

[31] Lupton JE, Lilley M, Butterfield D*, et al.* (2008) Venting of a separate CO_2-rich gas phase from submarine arc volcanoes; examples from the Mariana and Tonga-Kermadec Arcs. *Journal of Geophysical Research* **113**: B08S12.

[32] McClain CR (2007) Seamounts: identity crisis or split personality? *Journal of Biogeography* **34**: 2001-2008.

[33] Mottl MJ, Komor SC, Fryer P & Moyer CL (2003) Deep-slab fluids fuel extremophilic *Archaea* on a Mariana forearc serpentinite mud volcano: Ocean Drilling Program Leg 195. *Geochemistry Geophysics Geosystems* **4**: 1-14.

[34] Moussard H, Corre E, Cambon-Bonavita M-A, Fouquet Y & Jeanthon C (2006) Novel uncultured *Epsilonproteobacteria* dominate a filamentous sulphur mat from the 13 °N hydrothermal vent field, East Pacific Rise. *FEMS Microbiology Ecology* **58**: 449-463.

[35] Moyer CL, Dobbs FC & Karl DM (1994) Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi seamount, Hawaii. *Applied and Environmental Microbiology* **60**: 871-879.

[36] Moyer CL, Dobbs FC & Karl DM (1995) Phylogenetic diversity of the bacterial community from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. *Applied and Environmental Microbiology* **61**: 1555-1562.

[37] Nadkarni MA, Martin FE, Jacques NA & Hunter N (2002) Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* **148**: 257-266.

[38] Nakagawa S, Inagaki F, Takai K, Horikoshi K & Sako Y (2005) *Thioreductor micantisoli* gen. nov., sp. nov., a novel mesophilic, sulfur-reducing chemolithoautotroph within the ε–Proteobacteria isolated from hydrothermal sediments in the Mid-OkinawaTrough. *International Journal of Systematic and Evolutionary Microbiology* **55**: 599-605.

[39] Nakagawa S, Takai K, Inagaki F, Hirayama H, Nunoura T, Horikoshi K & Sako Y (2005) Distribution, phylogenetic diversity and physiological characteristics of epsilon-*Proteobacteria* in a deep-sea hydrothermal field. *Environmental Microbiology* **7**: 1619-1632.

[40] Nakagawa S, Takai K, Inagaki F*, et al.* (2005) Variability in microbial community and venting chemistry in a sediment-hosted backarc hydrothermal system: Impacts of subseafloor phase-separation. *FEMS Microbiology Ecology* **54**: 141-155.

[41] Opatkiewicz AD, Butterfield DA & Baross JA (2009) Individual hydrothermal vents at Axial Seamount harbor distinct subseafloor microbial communities. *FEMS Microbiology Ecology* **70**: 81-92.

[42] Papke RT & Ward DM (2004) The importance of physical isolation to microbial diversification. *FEMS Microbiology Ecology* **48**: 293-303.

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

[44] Quince C, Lanzen A, Curtis TP*, et al.* (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. *Nature Methods* **9**: 639-641.

[45] Rassa AC, McAllister SM, Safran SA & Moyer CL (2009) *Zeta-Proteobacteria* Dominate the Colonization and Formation of Microbial Mats in Low-Temperature Hydrothermal Vents at Loihi Seamount, Hawaii. *Geomicrobiology Journal* **26**: 623 - 638.

[46] Reeder J & Knight R (2009) The 'rare biosphere': a reality check. *Nature Methods* **6**: 636-637.

[47] Reno ML, Held NL, Fields CJ, Burke PV & Whitaker RJ (2009) Biogeography of the *Sulfolobus islandicus* pan-genome. *Proceedings of the National Academy of Sciences* **106**: 8605-8610.

[48] Resing JA, Baker ET, Lupton JE, Walker SL, Butterfield DA, Massoth GJ & Nakamura K (2009) Chemistry of hydrothermal plumes above submarine volcanoes of the Mariana Arc. *Geochemistry Geophysics Geosystems* **10**: 1-23.

[49] Reysenbach A-L, Longnecker K & Kirshtein J (2000) Novel bacterial and archaeal lineages from an in situ growth chamber deployed at a Mid-Atlantic Ridge hydrothermal vent. *Applied and Environmental Microbiology* **66**: 3798-3806.

[50] Samadi S, Bottan L, Macpherson E, De Forges B & Boisselier M-C (2006) Seamount endemism questioned by the geographic distribution and population genetic structure of marine invertebrates. *Marine Biology* **149**: 1463-1475.

[51] Schloss PD, Westcott SL, Ryabin T*, et al.* (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**: 7537-7541.

[52] Schrenk MO, Huber JA & Edwards KJ (2010) Microbial Provinces in the Subseafloor. *Annual Review of Marine Science* **2**: 279-304.

[53] Staudigel H, Hart SR, Pile A*, et al.* (2006) Vailulu'u Seamount, Samoa: Life and death on an active submarine volcano. *Proceedings of the National Academy of Sciences* **103**: 6448-6453.

[54] Stern RJ, Bloomer SH, Lin P-N, Ito E & Morris J (1988) Shoshonitic magmas in nascent arcs: New evidence from submarine volcanoes in the northern Marianas. *Geology* **16**: 426-430.

[55] Stocks KI & Hart PJ (2007) Biogeography and biodiversity of seamounts. *Seamounts: ecology, conservation and management,* Vol. 12 (Pitcher TJ, Morato T, Hart PJB, Clar MR, Haggan N & Sanots RS, ed.^eds.), p.^pp. 255-281. Blackwell, Oxford.

[56] Sun Y, Cai Y, Liu L, Yu F, Farrell ML, McKendree W & Farmerie W (2009) ESPRIT: estimating species richness using large collections of 16S rRNA pyrosequences. *Nucl. Acids Res.* **37**: e76-.

[57] Takacs-Vesbach C, Mitchell K, Jackson-Weaver O & Reysenbach A-L (2008) Volcanic calderas delineate biogeographic provinces among Yellowstone thermophiles. *Environmental Microbiology* **10**: 1681-1689.

[58] Takai K & Horikoshi K (2000) Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. *Applied and Environmental Microbiology* **66**: 5066-5072.

[59] Takai K, Hirayama H, Nakagawa T, Suzuki Y, Nealson KH & Horikoshi K (2005) *Lebetimonas acidiphila* gen. nov., sp. nov., a novel thermophilic, acidophilic, hydrogen-oxidizing chemolithoautotroph within the '*Epsilonproteobacteria*', isolated from a deep-sea hydrothermal fumarole in the Mariana Arc. *International Journal of Systematic and Evolutionary Microbiology* **55**: 183-189.

[60] Takai K, Campbell BJ, Cary SC*, et al.* (2005) Enzymatic and genetic characterization of carbon and energy metabolisms by deep-sea hydrothermal chemolithautorophic isolates of *Epsilonproteobacteria*. *Applied and Environmental Microbiology* **71**: 7310-7320.

[61] Telford RJ, Vandvik V & Birks HJB (2006) Dispersal limitations matter for microbial morphospecies. *Science* **312**: 1015.

[62] Tunnicliffe V, Davies KTA, Butterfield DA, Embley RW, Rose JM & Chadwick Jr WW (2009) Survival of mussels in extremely acidic waters on a submarine volcano. *Nature Geosciences* **2**: 344-348.

[63] Whitaker RJ (1998) *Island Biogeography*. Oxford University Press, Oxford.

[64] Whitaker RJ (2006) Allopatric origins of microbial species. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**: 1975-1984.

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

Table 1. Characteristics of samples used for analysis.

Sample	Seamount	Vent Name	Tmax (°C)	Tavg (°C)	Depth (m)	pH	Lat DecDeg	Long DecDeg	Volume (ml)	Cells ml-1 (±95% CI)
FS431	Forecast	Snail Scrum	16.0	6.0	1448	6.56	13.395	143.920	2050	3.38×10^5 (3.92×10^4)
FS432	Forecast	Homer Vent	19.0	6.5	1451	6.23	13.395	143.920	2015	2.37×10^5 (2.70×10^4)
FS445	NW Rota-1	Brimstone	27.9	19.7	560	1.64	14.601	144.775	3004	no data
FS446	NW Rota-1	Iceberg	53.8	48.0	534	3.75	14.601	144.776	3007	8.54×10^4 (8.84×10^3)
FS447	NW Rota-1	Sandy Saddle	35.6	29.0	521	2.62	14.601	144.776	1960	1.23×10^5 (1.78×10^4)
FS448	NW Rota-1	Fault Shrimp	25.9	25.0	584	5.25	14.601	144.777	3010	3.49×10^5 (2.62×10^4)
FS449	NW Rota-1	Scarp Top	17.6	15.1	568	5.09	14.601	144.778	2000	3.07×10^5 (3.17×10^4)
FS467	NW Eifuku	Sulfur Flow	52.7	42.9	1612	5.26	21.487	144.042	4000	1.34×10^6 (1.16×10^5)
FS468	NW Eifuku	Cliff House	54.1	45.1	1578	5.28	21.487	144.042	5001	1.95×10^6 (1.64×10^5)
FS473	Daikoku	NE Pit	15.6	15.3	438	5.89	21.325	144.193	2019	2.17×10^5 (1.46×10^4)
FS475	Daikoku	Alka Seltzer	64.5	45.5	414	3.36	21.325	144.191	1317	1.67×10^5 (2.40×10^4)
FS479	Nikko	North Vent	96.1	80.2	458	5.54	23.081	142.325	4000	4.18×10^4 (7.33×10^3)
FS480	Nikko	Tubeworm Hangover	26.1	24.1	445	5.05	23.079	142.326	3522	4.95×10^5 (3.56×10^4)
FS481	Nikko	Top Vent	35.4	32.6	413	6.21	23.080	142.327	3059	2.11×10^5 (6.21×10^4)

Table 2. Sequencing statistics and qPCR results for each sample

	Total number of Bacterial V6 Tag Sequences[a]	Total number of *epsilonproteobacteria* V6 Tag Sequences[a]	Percent *epsilonproteobacteria* V6 Tag Sequences	Percent Bacteria[b]	Percent Archaea[b]
FS431	15219	3930	25.8%	98.33 (± 0.26)	1.67 (± 0.26)
FS432	16393	4248	25.9%	98.61 (± 0.16)	1.39 (± 0.16)
FS445	8411	1464	17.4%	92.04 (+ 1.21)	7.96 (+ 1.21)
FS446	12478	8684	69.6%	97.31 (± 0.07)	2.69 (± 0.07)
FS447	9982	2515	25.2%	97.02 (± 0.42)	2.98 (± 0.42)
FS448	19343	3033	15.7%	96.79 (+ 0.08)	3.21 (+ 0.08)
FS449	11707	3915	33.4%	97.64 (± 0.24)	2.36 (± 0.24)
FS467	12279	5498	44.8%	98.80 (± 0.19)	1.20 (± 0.19)
FS468	62521	17283	27.6%	98.54 (+ 0.46)	1.46 (+ 0.46)
FS473	21459	786	3.7%	97.55 (± 0.04)	2.45 (± 0.04)
FS475	10804	8493	78.6%	99.53 (± 0.06)	0.47 (± 0.06)
FS479	11890	3063	25.8%	97.51 (+ 0.21)	2.49 (+ 0.21)
FS480	9255	8093	87.4%	99.46 (± 0.00)	0.54 (± 0.00)
FS481	8946	3762	42.1%	99.32 (± 0.00)	0.68 (± 0.00)

[a] Trimmed reads that passed quality control as described in Methods

[b] Based on qPCR as described in methods

Table 3. Pairwise R values from ANOSIM test between groupings and seamounts

	Forecast	NW Rota-1	NW Eifuku	Daikoku	Nikko
Forecast					
NW Rota-1	0.95				
NW Eifuku	1.00	0.64			
Daikoku	1.00	0.73	0.50		
Nikko	0.67	0.82	0.08	-0.33	

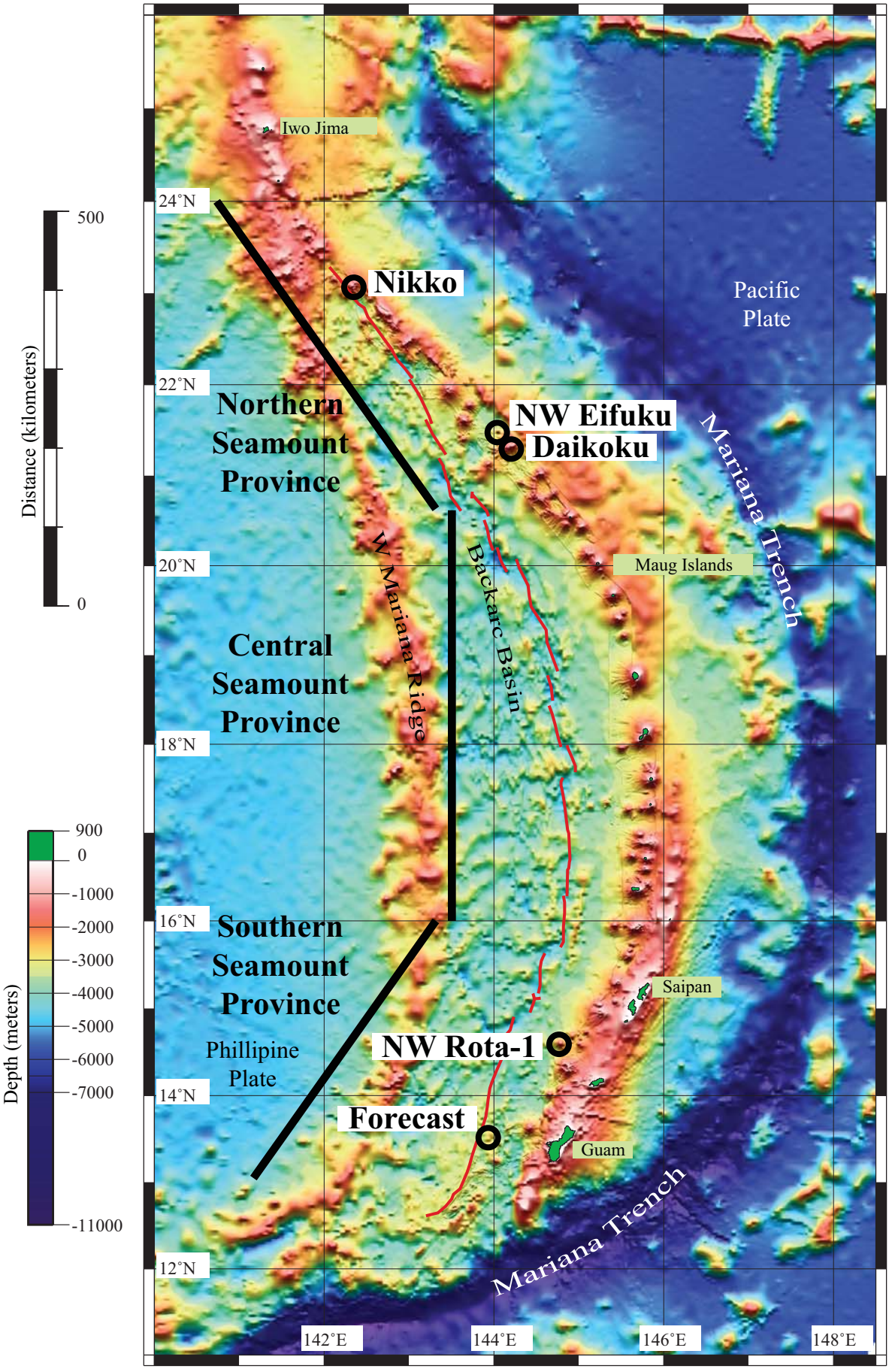
Figure 1. Map of the Mariana Arc and five volcanoes where vent fluids were collected. Map courtesy of R. Embley and S. Merle, NOAA/PMEL.

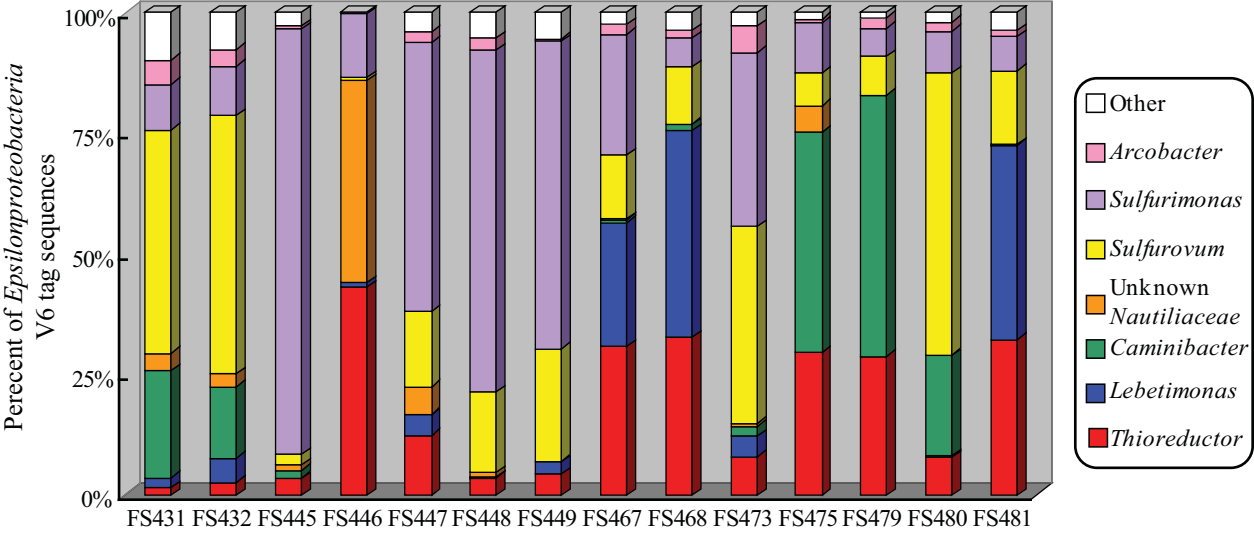
Figure 2. Taxonomic breakdown and relative abundance at the genus level for the individual tag sequences of *Epsilonbacteri*a in each vent fluid sample. Only those genera that occurred more than 1000 times across all datasets are labeled. All others are lumped into "Other."

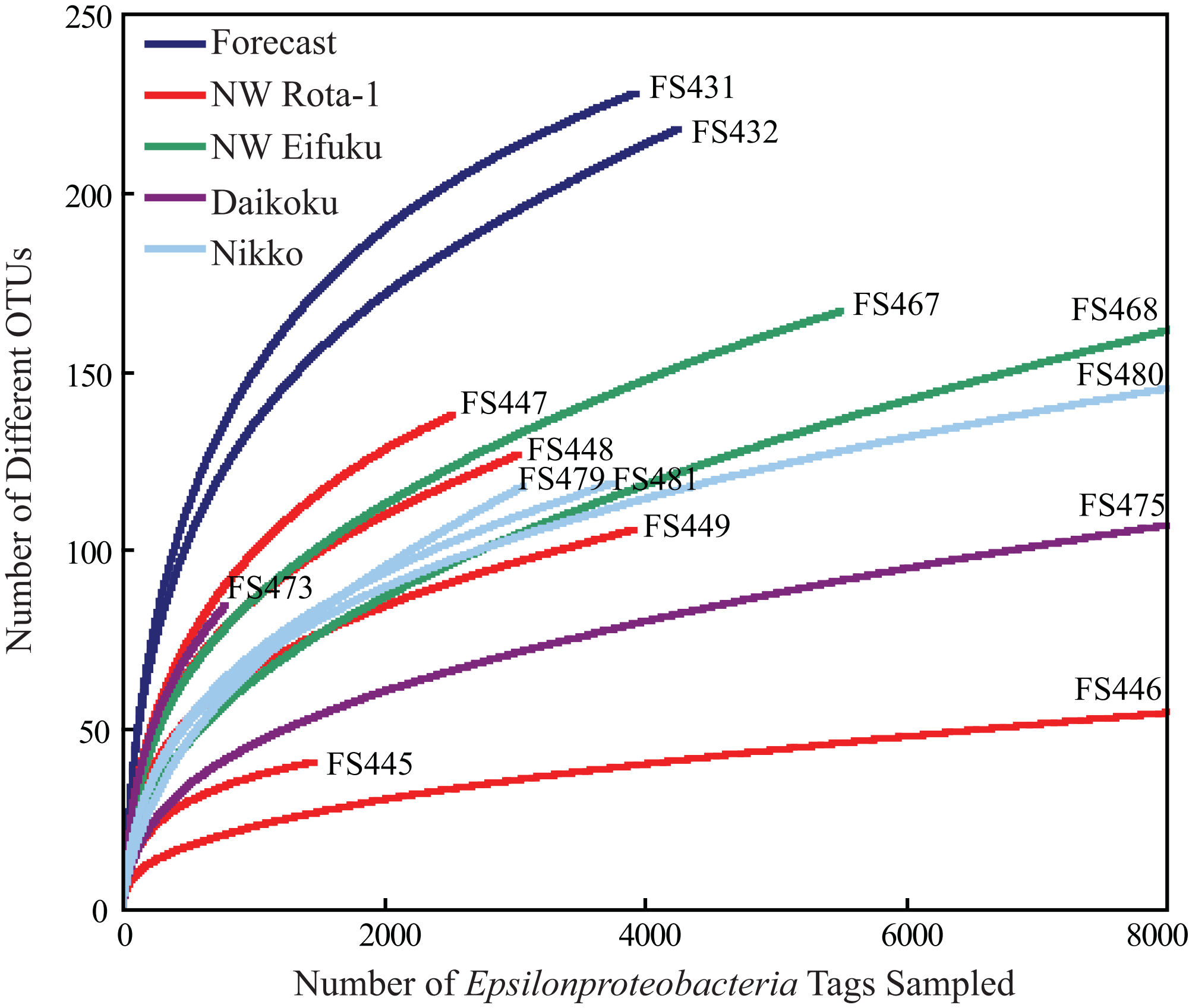
Figure 3. Rarefaction curves at the 3% difference level for each of the vent fluid samples.

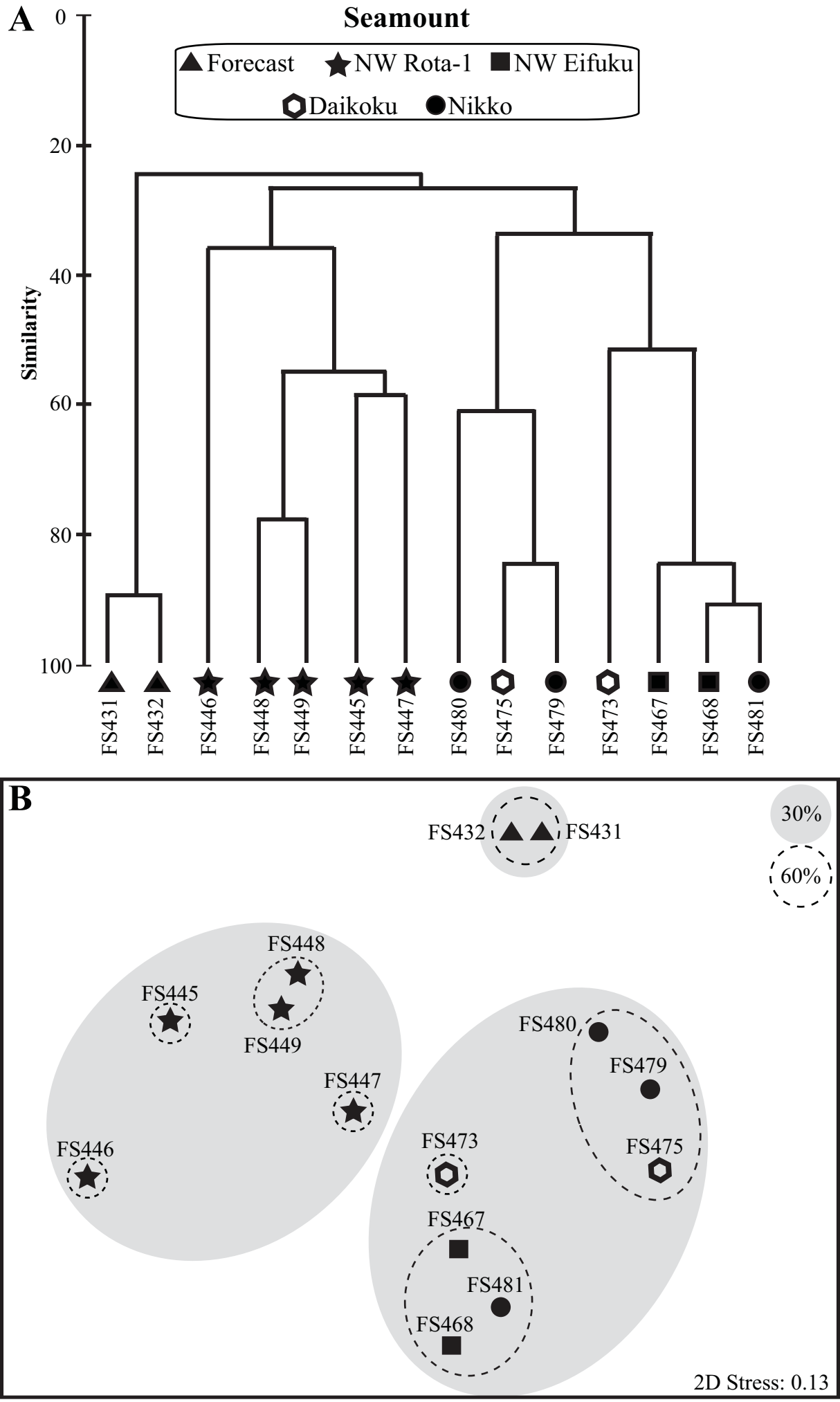
Figure 4. (A) Cluster diagram and (B) nMDS 2D similarity plot comparing vent samples at the 3% difference level of *Epsilonproteobacteria* with each sample labeled according to seamount. The circles represent similarity boundaries from the cluster diagram.

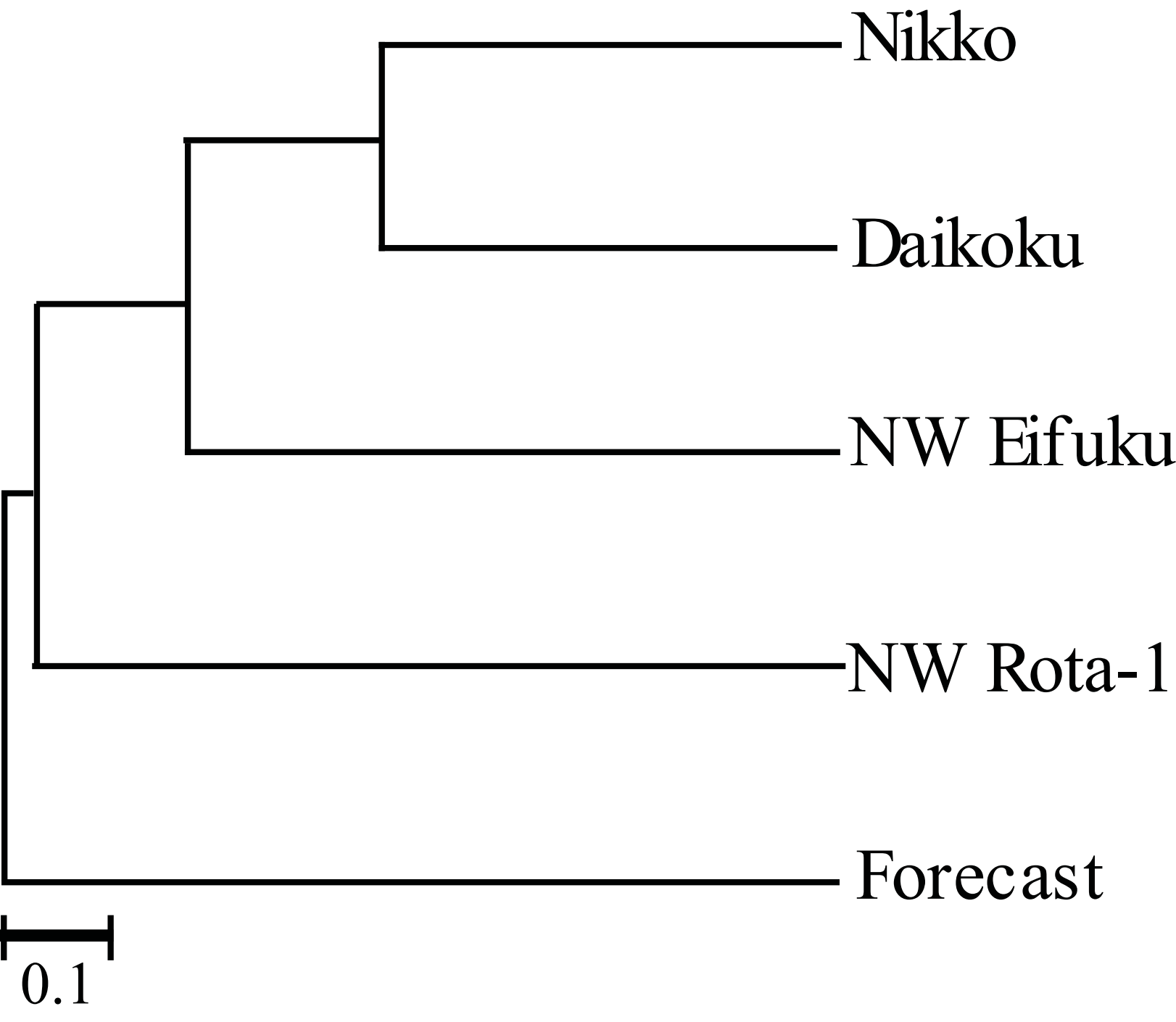
Figure 5. Similarity of pooled samples at the 3% difference level from each seamount based on the Morisita-Horn measure. The length of the reference bar represents a distance of 0.10.


Iwo Jima
24°N
Nikko
Pacific Plate
22°N
Northern Seamount Province
NW Eifuku
Daikoku
Mariana Trench
W Mariana Ridge
Backarc Basin
Maug Islands
20°N
Central Seamount Province
18°N
16°N
Southern Seamount Province
Saipan
NW Rota-1
Phillipine Plate
14°N
Forecast
Guam
Mariana Trench
12°N
142°E
144°E
146°E
148°E
Distance (kilometers)
500
0
Depth (meters)
900
0
-1000
-2000
-3000
-4000
-5000
-6000
-7000
-11000


Perecent of Epsilonproteobacteria V6 tag sequences
100%
75%
50%
25%
0%
FS431 FS432 FS445 FS446 FS447 FS448 FS449 FS467 FS468 FS473 FS475 FS479 FS480 FS481
Other
Arcobacter
Sulfurimonas
Sulfurovum
Unknown Nautiliaceae
Caminibacter
Lebetimonas
Thioreductor


Forecast
NW Rota-1
NW Eifuku
Daikoku
Nikko
FS431
FS432
FS467
FS468
FS480
FS447
FS448
FS479 FS481
FS475
FS449
FS473
FS446
FS445
250
200
150
100
50
0
0
2000
4000
6000
8000
Number of Different OTUs
Number of *Epsilonproteobacteria* Tags Sampled


A
Seamount
Forecast
NW Rota-1
NW Eifuku
Daikoku
Nikko
Similarity
0
20
40
60
80
100
FS431
FS432
FS446
FS448
FS449
FS445
FS447
FS480
FS475
FS479
FS473
FS467
FS468
FS481
B
30%
60%
FS432
FS431
FS448
FS445
FS449
FS447
FS446
FS480
FS479
FS475
FS473
FS467
FS481
FS468
2D Stress: 0.13


Nikko
Daikoku
NW Eifuku
NW Rota-1
Forecast
0.1