

# An interlaboratory comparison of 16S rRNA gene-based terminal restriction fragment length polymorphism and sequencing methods for assessing microbial diversity of seafloor basalts

Beth Orcutt,<sup>1</sup> Brad Bailey,<sup>2</sup> Hubert Staudigel,<sup>2</sup>  
Bradley M. Tebo<sup>3</sup> and Katrina J. Edwards<sup>1\*</sup>

<sup>1</sup>Geomicrobiology Group, Department of Biological Sciences, Marine Environmental Biology, University of Southern California, Los Angeles, CA 90089, USA.

<sup>2</sup>Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, La Jolla, CA 92037, USA.

<sup>3</sup>Division of Environmental and Biomolecular Systems, Oregon Health and Science University, Beaverton, OR 97006, USA.

## Summary

**We present an interlaboratory comparison between full-length 16S rRNA gene sequence analysis and terminal restriction fragment length polymorphism (TRFLP) for microbial communities hosted on seafloor basaltic lavas, with the goal of evaluating how similarly these two different DNA-based methods used in two independent labs would estimate the microbial diversity of the same basalt samples. Two samples were selected for these analyses based on differences detected in the overall levels of microbial diversity between them. Richness estimators indicate that TRFLP analysis significantly underestimates the richness of the relatively high-diversity seafloor basalt microbial community: at least 50% of species from the high-diversity site are missed by TRFLP. However, both methods reveal similar dominant species from the samples, and they predict similar levels of relative diversity between the two samples. Importantly, these results suggest that DNA-extraction or PCR-related bias between the two laboratories is minimal. We conclude that TRFLP may be useful for relative comparisons of diversity between basalt samples, for identifying dominant species, and for estimating the**

**richness and evenness of low-diversity, skewed populations of seafloor basalt microbial communities, but that TRFLP may miss a majority of species in relatively highly diverse samples.**

## Introduction

Application of the proper metrics for assessing, measuring and quantifying microbial populations in complex natural systems is a current and long-standing challenge in microbial diversity studies (reviewed in Hughes *et al.*, 2001). Community genetic fingerprinting methods, such as terminal restriction fragment length polymorphism (TRFLP; Liu *et al.*, 1997), have been used repeatedly to investigate comparative community composition owing to the procedures being relatively rapid, high-throughput and inexpensive. For highly diverse microbial communities, however, there are conflicting results whether different fingerprinting methods determine similar (e.g. Hartmann *et al.*, 2005) or different (e.g. Danovaro *et al.*, 2006) levels of diversity, although part of the disagreement between methodological approaches may be due to focusing in some cases on the coding 16S rRNA gene and in others on the more highly variable, non-coding intergenic regions of the DNA sequence. In addition, the utility of fingerprinting methods to accurately calculate community diversity in complex communities has been criticized due to inability of the methods to detect low-abundance (< 1% of the community) and rare taxa (Dunbar *et al.*, 2000; Blackwood *et al.*, 2003; Engebretson and Moyer, 2003; Bent *et al.*, 2007). Recent innovations, such as pyrosequencing, now suggest that microbial communities contain even greater diversity than previously accepted (Sogin *et al.*, 2006; Huber *et al.*, 2007). It is unknown how the rare members of any ecosystem translate to ecological or biogeochemical function, although it is suggested that the rare taxa may serve as a reservoir of genomic innovation to allow microbial communities to adapt to changing environmental conditions (Sogin *et al.*, 2006). If fingerprinting methods exclude rare taxa, are they useful for predicting metabolic potential in a microbial ecosystem?

A further challenge for microbial ecologists is the reproducibility of microbial biodiversity surveys on similar

Received 17 July, 2008; accepted 31 January, 2009. \*For correspondence. E-mail kje@usc.edu; Tel. (+1) 213 821 4390; Fax (+1) 213 740 8123.

Re-use of this article is permitted in accordance with the Creative Commons Deed, Attribution 2.5, which does not permit commercial exploitation.

samples using different methods. As one example, surveys conducted on similar deep marine subsurface sediments using different methods have resulted in contradictory interpretations of whether the domain *Bacteria* or *Archaea* dominate the deep biosphere microbial community (Biddle *et al.*, 2006; Inagaki *et al.*, 2006; Schippers and Neretin, 2006; Sørensen and Teske, 2006; Lipp *et al.*, 2008). In order to assess global and local biodiversity patterns and to correctly correlate microbial community form with function, it is essential that methods provide predictable and ecologically meaningful information, even if biased (Hughes Martiny *et al.*, 2006).

In the present study, two independent laboratories utilized different approaches that target the 16S rRNA gene to evaluate diversity of seafloor basalt samples, and the results were cross-compared for the purposes of evaluating methods and interlaboratory biases in biodiversity assessments. Previous studies have shown that seafloor basalts harbour some of the most diverse microbial communities on Earth (Santelli *et al.*, 2008). Here, a set of two seafloor basalt samples collected from the Lō'ihi Seamount were analysed independently by both TRFLP analysis and 16S rRNA gene clone libraries as well as *in silico* TRFLP analysis of the 16S rRNA sequences. Commonly used diversity indices were applied to the results of both the community fingerprinting analysis by TRFLP as well as to the similarity matrices calculated from the gene sequences. As compared with previous computer-simulated comparisons between TRFLP patterns and 16S rRNA gene sequences (Liu *et al.*, 1997; Engebretson and Moyer, 2003; Blackwood *et al.*, 2007) and comparisons between restriction digests of 16S rRNA gene clones amplified from soils (Dunbar *et al.*, 1999; 2000), this study evaluated empirically derived diversity assessments generated independently by two labs from the same samples.

## Results

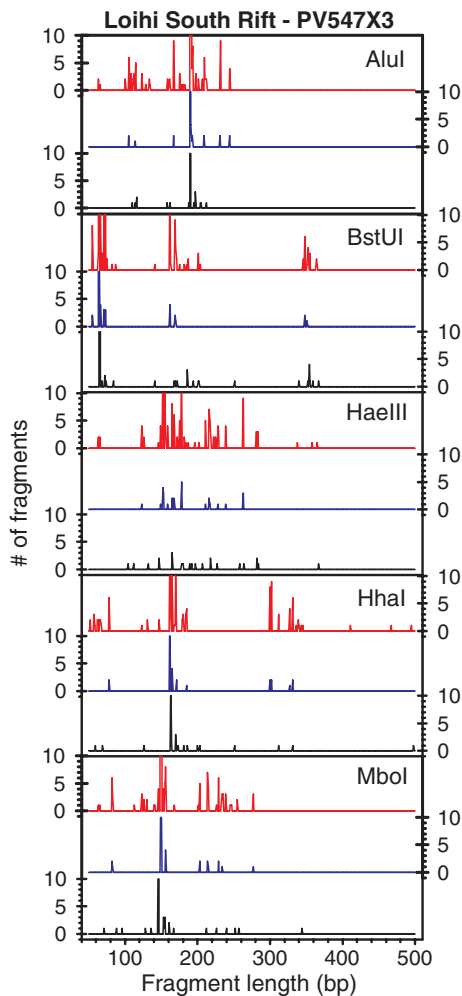
Two seafloor basalt samples were the focus of this comparative study, which had recently been identified as representing relatively high- and low-diversity end members among a suite of basalt samples from Hawai'i (Santelli *et al.*, 2008). The 'high diversity' sample originated from the Pisces Peak location on the Lō'ihi Seamount off the south coast of the big island of Hawai'i (sample PV549X2). The 'low diversity' end member was collected from the South Rift location on the southern flank of Lō'ihi (sample PV547X3).

From two subsamples of the same rocks, 16S rRNA gene clone libraries were constructed following DNA extraction in one lab, while TRFLP analysis was performed with DNA extracted in another lab. The two labs used different DNA extraction protocols and slightly different primers for PCR amplification (see *Experimental*

*procedures* below). From the clone library data, which contained 246 sequences from the high-diversity sample and 71 sequences from the low-diversity sample, the microbial diversity and richness were estimated from sequence alignments using the program DOTUR (Schloss and Handelsman, 2005) at a  $\geq 97\%$  sequence similarity definition for species, as described elsewhere (Santelli *et al.*, 2008). To allow comparison of the clone library data with the measured TRFLP analyses (see below), an *in silico* terminal restriction digestion (Ricke *et al.*, 2005) was performed on the clone library sequences to generate theoretical fragment patterns that could be directly compared with the measured TRFLP patterns. Figures 1 and 2 present the measured and predicted TRFLP patterns from the South Rift and Pisces Peak samples respectively.

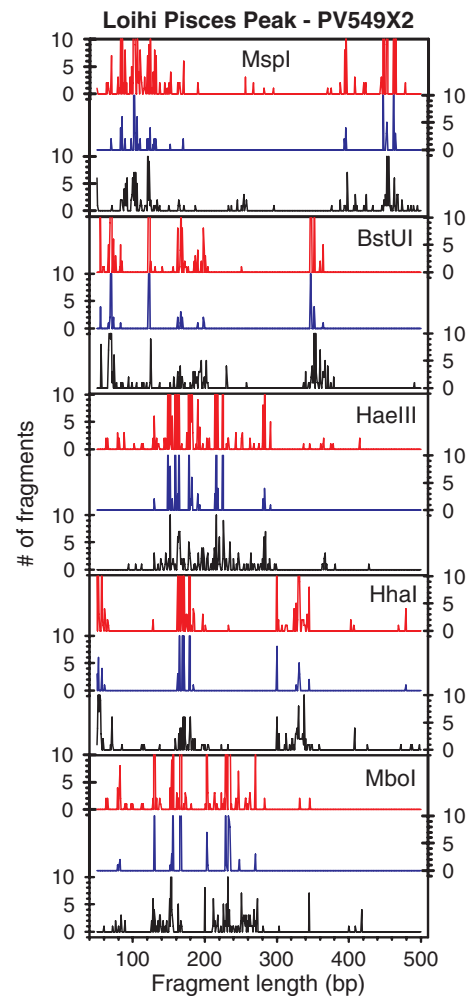
To evaluate how each method predicted microbial community diversity, a suite of classical non-parametric richness and evenness indices were calculated from the observed operational taxonomic unit (OTU) distribution generated by each of the sampling methods (Table 1). As shown previously (Santelli *et al.*, 2008), diversity and richness estimates based on clone library sequence alignments via DOTUR indicated that the Pisces Peak sample had higher richness with roughly 400 species (from abundance-based coverage (ACE) species richness estimator and Chao1 species richness estimator) as compared with the South Rift basalt, which was estimated to support approximately 100 species (from ACE and Chao1). Calculation of the Shannon–Weaver index ( $H'$ ) and the Simpson index ( $D$ ) from the sequence alignments also indicated that the Pisces Peak sample had a higher diversity than the South Rift basalt (respectively, diversity index of 4.77 versus 1.83 for  $H'$  and 0.0071 versus 0.38 for  $D$ ). In addition, the Shannon evenness index ( $J'$ ) indicated that the Pisces Peak sample had a more even microbial population (i.e. index closer to 1) than the South Rift sample.

When calculating species richness from the predicted *in silico* TRFLP OTU distributions (*is*TRFLP, Table 1), the ACE and Chao1 estimates ( $86 \pm 19$  and  $91 \pm 31$ , respectively, with the variation attributable to the differences between restriction enzyme considered) were similar to the sequence-derived values (roughly 100) for the low-diversity sample from South Rift, but the predicted species richness for the high-diversity sample from Pisces Peak was significantly lower, with a predicted richness of roughly 170 species as compared with ~400. Additionally, the Shannon  $H'$  and Simpson richness indices estimated lower richness from the *in silico* TRFLP OTU distributions for the Pisces Peak sample, but similar values were calculated for the South Rift sample by both methods. Notably, the calculation of species evenness ( $J'$ ) was similar by either method for both samples.



**Fig. 1.** Comparison of TRFLP patterns from the South Rift sample. Red lines indicate number of fragments from the measured TRFLP with the respective enzymes when assuming a threshold cut-off of 15 fluorescence units; blue lines indicate the number of fragments from the measured TRFLP with the respective enzymes when assuming a threshold cut-off of 50 fluorescence units; and black lines indicate the number of TRFLP fragments generated from *in silico* digest of nearly full-length 16S rRNA sequences using the respective enzyme pre-sets in the tRF-cut program in ARB. Some peaks are larger than y-scale given, small y-scales given to highlight rare OTUs (shorter peaks).

The prediction of species richness from the ACE and Chao1 estimators from the measured TRFLP analysis was significantly lower than from the other methods for both samples, regardless of whether a threshold cut-off value of 15 or 50 fluorescence units was assumed for converting TRFLP electropherogram data into species abundance patterns (see *Experimental procedures* section). For example, there were a predicted  $76 \pm 14$  species in the Pisces Peak basalt and  $47 \pm 8$  species in the South Rift basalt when assuming a 15 fluorescence unit cut-off value (TRFLP<sub>15</sub>, Table 1); even lower estimates of  $25 \pm 2$  species in the Pisces Peak sample and



**Fig. 2.** Comparison of TRFLP patterns from the Pisces Peak sample. Red lines indicate number of fragments from the measured TRFLP with the respective enzymes when assuming a threshold cut-off of 15 fluorescence units; blue lines indicate the number of fragments from the measured TRFLP with the respective enzymes when assuming a threshold cut-off of 50 fluorescence units; and black lines indicate the number of TRFLP fragments generated from *in silico* digest of nearly full-length 16S rRNA sequences using the respective enzyme pre-sets in the tRF-cut program in ARB. Some peaks are larger than y-scale given, small y-scales given to highlight rare OTUs (shorter peaks).

$13 \pm 4$  species in the South Rift sample were calculated when assuming a threshold value of 50 fluorescence units (TRFLP<sub>50</sub>, Table 1). The Shannon and Simpson richness indices also estimated significantly lower diversity for the Pisces Peak sample from the measured TRFLP patterns than from the other methods; however, the species evenness prediction ( $J'$ ) was similar but slightly lower. In contrast, for the lower diversity South Rift sample, the Shannon and Simpson richness indices estimated from the measured TRFLP data predicted a similar or slightly more diverse community than estimated using the other sequence based and *is*TRFLP methods. Although the

**Table 1.** Diversity and evenness indices calculated from various methods for the two seafloor basalt samples considered in this study.

| Sample      | Measure                          | ACE <sup>a</sup> | Chao <sup>b</sup> | Shannon $H'$ <sup>c</sup> | Shannon $J'$ <sup>d</sup> | Simpson <sup>e</sup> | $S_{obs}$ <sup>f</sup> | $n_1$ <sup>g</sup> | $n_2$ <sup>h</sup> |
|-------------|----------------------------------|------------------|-------------------|---------------------------|---------------------------|----------------------|------------------------|--------------------|--------------------|
| Pisces Peak | DOTUR <sup>i</sup>               | 404              | 416               | 4.77                      | 0.95                      | 0.0071               | 148                    | 104                | 19                 |
|             | <i>is</i> TRFLP <sup>j</sup>     | 169 ± 26         | 168 ± 25          | 4.16 ± 0.2                | 0.91 ± 0.02               | 0.020 ± 0.08         | 96 ± 11                | 50 ± 8             | 17 ± 4             |
|             | TRFLP <sub>15</sub> <sup>k</sup> | 76 ± 14          | 69 ± 10           | 3.08 ± 0.2                | 0.76 ± 0.04               | 0.08 ± 0.02          | 59 ± 10                | 17 ± 4             | 15 ± 6             |
|             | TRFLP <sub>50</sub> <sup>l</sup> | 25 ± 2           | 23 ± 2            | 2.55 ± 0.2                | 0.84 ± 0.04               | 0.10 ± 0.03          | 21 ± 3                 | 4 ± 1              | 4 ± 4              |
| South Rift  | DOTUR <sup>i</sup>               | 110              | 87.3              | 1.83                      | 0.58                      | 0.38                 | 24                     | 20                 | 2                  |
|             | <i>is</i> TRFLP <sup>j</sup>     | 86 ± 19          | 91 ± 31           | 1.92 ± 0.2                | 0.62 ± 0.05               | 0.32 ± 0.08          | 22 ± 2                 | 17 ± 2             | 2 ± 1              |
|             | TRFLP <sub>15</sub> <sup>k</sup> | 47 ± 8           | 43 ± 7            | 2.50 ± 0.5                | 0.70 ± 0.13               | 0.17 ± 0.11          | 35 ± 4                 | 11 ± 3             | 9 ± 2              |
|             | TRFLP <sub>50</sub> <sup>l</sup> | 13 ± 4           | 13 ± 4            | 1.81 ± 0.4                | 0.75 ± 0.13               | 0.26 ± 0.15          | 11 ± 2                 | 4 ± 2              | 4 ± 1              |

a. ACE, predicted species richness using abundance-based coverage estimator, calculated using formulation from Chao and Lee (1992).

b. Chao, predicted species richness using Chao1 estimator, calculated using formulation from Chao and colleagues (1992).

c. Shannon  $H'$ , Shannon–Weaver index of diversity, higher numbers indicate higher diversity, calculated using formulation from Schloss and Handelsman (2005).

d. Shannon  $J'$ , Shannon evenness, values can range from 0 to 1, with 1 representing perfect evenness, calculated using formulation from Schloss and Handelsman (2005).

e. Simpson, Simpson index of diversity, values can range from 0 to 1 with higher values indicating a skewed population with dominant members, calculated using formulation from Schloss and Handelsman (2005).

f.  $S_{obs}$ , observed OTUs.

g. Number of OTUs with only one member.

h. Number of OTUs with only two members.

i. DOTUR, based on neighbor-joining alignment at 97% sequence similarity of nearly full-length 16S rRNA sequences in ARB following application of a filter for bacteria for the positions 1218–42590 and the Jukes Cantor correction.

j. *is*TRFLP: estimated using all of the TRFLP fragments generated from the *in silico* digest of the sequences recovered from 16S rRNA clone library; average of five digests with different restriction enzymes ± standard deviation of the average.

k. TRFLP<sub>15</sub>: estimated from the actual TRFLP fragments assuming a threshold cut-off value of 15 fluorescence units; average of five digests with different restriction enzymes ± standard deviation of the average.

l. TRFLP<sub>50</sub>: estimated from the actual TRFLP fragments assuming a threshold cut-off value of 50 fluorescence units; average of five digests with different restriction enzymes ± standard deviation of the average.

OTU sample sizes ( $S_{obs}$ , Table 1) for the low-diversity sample are relatively small, they are still expected to be significant enough that the ACE and Chao1 richness estimators are not overly biased by sample size [i.e. while some sample size bias on the predicted richness estimators is expected (see Youssef and Elshahed, 2008),  $S_{obs}$  in our samples is greater than the square root of twice the estimate of richness; Colwell and Coddington, 1994]. Notably, all methods calculate that the Pisces Peak sample harbours a more diverse microbial community than the South Rift sample.

A peak-by-peak correlation between the measured TRFLP fragment peaks and those predicted from the *in silico* digestion of the clone library sequences (Figs 1 and

2, Table 2) reveals that peaks predicted from the *in silico* TRFLP digestion were correlated perfectly or within 1–2 base pairs (bp) to 37% and 72% of the peaks from the measured TRFLP<sub>15</sub> patterns for the South Rift and Pisces Peak samples respectively (the numbers shift to 45% and 88% of peaks respectively, when using TRFLP<sub>50</sub>). The viability of correlating peaks offset by 1–2 bp is validated by the variation between true and observed fragments based on the relative composition of the DNA sequences (Kaplan and Kitts, 2003). When weighted by frequency of appearance of a particular fragment, this corresponds to 67% and 92%, respectively, of PCR amplicons from the measured TRFLP<sub>15</sub> analysis being matched by fragments in the *in silico* TRFLP analysis (or 75% and 93% when

**Table 2.** Peak-to-peak correlation (as a percentage of total OTUs) between measured TRFLP patterns and predicted TRFLP patterns from an *in silico* digestion of 16S rRNA clone library data.

|                                      | South Rift      |                 | Pisces Peak     |                 |
|--------------------------------------|-----------------|-----------------|-----------------|-----------------|
|                                      | 15 f.u. cut-off | 50 f.u. cut-off | 15 f.u. cut-off | 50 f.u. cut-off |
| TRFLP peaks perfect match            | 12 ± 4          | 8 ± 10          | 42 ± 4          | 56 ± 11         |
| TRFLP peaks off by 1–2 bp            | 25 ± 13         | 37 ± 29         | 30 ± 4          | 32 ± 12         |
| TRFLP peaks not matched              | 63 ± 12         | 55 ± 20         | 28 ± 3          | 12 ± 7          |
| TRFLP 'members' identified           | 67 ± 11         | 75 ± 9          | 92 ± 3          | 93 ± 3          |
| <i>in silico</i> peaks perfect match | 24 ± 13         | 6 ± 8           | 29 ± 4          | 15 ± 4          |
| <i>in silico</i> peaks off by 1–2 bp | 45 ± 15         | 20 ± 15         | 21 ± 4          | 8 ± 3           |
| <i>in silico</i> peaks not matched   | 30 ± 7          | 77 ± 3          | 50 ± 6          | 77 ± 3          |

Data are presented for the correlations when the measured TRFLP fragments are assumed versus both 15- and 50-fluorescence-unit (f.u.) threshold cut-off values.

using TRFLP<sub>50</sub>). While the dominant species were recovered in both analyses, a larger proportion of rare species was recovered in the clone library data. Roughly one-third to two-thirds of the TRFLP peaks predicted from the *in silico* digestion are not matched in the measured TRFLP<sub>15</sub> or TRFLP<sub>50</sub> analyses.

## Discussion

Seafloor basalts harbour some of the most diverse microbial communities on Earth (Santelli *et al.*, 2008). Previous analysis indicates that the two samples used in this study represent both relatively low (Lô'ihi South Rift basalt) and high (Lô'ihi Pisces Peak basalt) diversity end members (Santelli *et al.*, 2008). A major goal of the present study was to evaluate how similarly two different DNA-based methods (i.e. TRFLP and 16S rRNA gene clone libraries) conducted in two independent labs would estimate the microbial diversity of these basalt samples. Although there are differences between the two methodological approaches to examining diversity of seafloor basalts (i.e. different DNA extraction methods and PCR conditions, see *Experimental procedures* below), it is notable that both methods retrieved similar dominant species from the samples, and that they predicted similar levels of relative diversity between the two samples. For instance, as shown in Fig. 1, the major species (i.e. largest peak) that dominated the low-diversity South Rift sample was retrieved by both methods, regardless of the enzyme considered for TRFLP. Similarly, the measured and *in silico* TRFLP patterns from the Pisces Peak sample are qualitatively similar, with dominant peaks often matching between the two methods (Fig. 2). Although predicting a much lower richness level, the calculation of richness from the measured TRFLP analyses also indicates that the Pisces Peak basalt microbial community is more diverse than the South Rift basalt community (Table 1). Importantly, these results suggest that there was minimal apparent DNA-extraction or PCR-related bias between the two laboratories although different protocols were used. Although, theoretically, it is not unexpected that the two approaches would uncover similar communities, this is rarely evaluated in molecular diversity studies and is often a considerable uncertainty in making cross-comparisons between studies.

The comparison of richness estimators calculated from the two data sets reveals that TRFLP analysis significantly underestimated the richness of the relatively high-diversity seafloor basalt microbial community (Table 1). Depending on the richness estimator considered, richness predicted by TRFLP analysis for the low-diversity sample was roughly the same (using Shannon and Simpson indices) or lower (ACE and Chao1 estimators) than the richness predicted by sequence similarity.

Although the richness estimates were different between the two methodologies, the predictions of community evenness were similar. It should be emphasized that the diversity indices calculated based on 16S rRNA gene clone similarity used a definition of  $\geq 97\%$  similar sequence for determining the number of species. If a higher threshold had been used (i.e. 99% sequence similarity), the number of species predicted from the clone library data would have been even greater. The  $\geq 97\%$  cut-off is a commonly recognized level for comparative analysis in environmental microbial communities. Nonetheless, it is likely that at least 50% of species from the Pisces Peak samples are missed by the laboratory-based TRFLP analysis (Table 2, assuming no bias from DNA extraction of PCR primer conditions), as parametric-based rarefaction analysis of the sequence similarity data for this sample (from Santelli *et al.*, 2008), which is a measure of the number of species observed per sampling effort, indicates that significantly more sampling would be needed to measure the full diversity of the sample.

There are at least two reasons why the richness estimated by TRFLP is lower than that estimated from clone library sequence alignment, as suggested previously for simulated comparisons of TRFLP and clone sequences (Engebretson and Moyer, 2003; Blackwood *et al.*, 2007). The first explanation is that TRFLP misses the majority of rare species, likely due to detection limits for resolving unique sequence variants. For example, if the 16S rRNA gene from a rare species did not generate enough fluorescently labelled PCR amplicons, this species may be missed by the limit of detection for the TRFLP analysis. An evaluation of how many rare species are captured by each method can be seen in the number of species that were observed only once ( $n_1$ ) or twice ( $n_2$ ) using the various methods (Table 1). For instance, in the clone libraries, 104 and 20 different species were observed only once in the Pisces Peak and South Rift samples, respectively, whereas only 17 and 11 species, respectively, were observed only once in the measured TRFLP<sub>15</sub> analysis, and only four species were observed once in each sample in the TRFLP<sub>50</sub> analysis. This trend is also evident in Figs 1 and 2, where more small peaks appear in the *in silico* TRFLP patterns than in the measured TRFLP<sub>15</sub> or TRFLP<sub>50</sub> patterns.

The second explanation for the lower estimations of richness from TRFLP as compared with the clone library data is the occurrence of binning in TRFLP, where two different species are counted as one TRFLP OTU because they generate the same size fragment. This can be shown empirically in the *in silico* TRFLP analysis of the clone library data, especially for the higher diversity Pisces Peak sample. For one example, the peak seen in the *in silico* *Bst*UI digestion of the Pisces Peak sample at a fragment size of 70 bp (Fig. 2) contains predicted frag-

ments derived from a diverse range of species from the phyla *Bacterioidetes*, *Firmicutes* and *Planctomycetes* as well as the Alpha-, Delta-, and Gamma- divisions of the phyla *Proteobacteria* (data not shown). *Bst*UI digestion generates 70 bp fragments from 14 different species as defined by gene sequence similarity, which in the predicted TRFLP analysis would be binned as one peak, or species. This result of multiple species binning into one OTU has been demonstrated before (e.g. Clement *et al.*, 1998). In the high-diversity sample from Pisces Peak, the degree of binning is evident in the difference between the richness estimators calculated from the sequence alignment and the *in silico* digestions (DOTUR versus *is*TRFLP in Table 1) – the *in silico*-based calculations predict nearly two-thirds less richness when estimated by the ACE and Chao1 estimators. For comparison, in randomly simulated microbial communities with richness levels of 100 species, the restriction enzymes considered in this study would correctly assign 60% or less of the expected OTUs (Engebretson and Moyer, 2003). Both of the above trends of binning and exclusion of rare taxa are also evident in a comparison of diversity indices calculated by TRFLP and restriction digestion of 16S rRNA gene clones in soils (Dunbar *et al.*, 1999; 2000). Binning is evident regardless of which restriction enzyme is considered. In principal, though, it is possible to resolve specific instances of binning in individual data sets by employing multiple restriction enzymes, as sequences that bin together with one restriction enzyme might produce different-sized fragments when targeted by a different restriction enzyme.

Results of these studies indicate that TRFLP could be useful for relative comparisons of diversity between samples, for identifying dominant species and for estimating the richness and evenness of low-diversity, skewed populations of seafloor basalt microbial communities. This observation supports the claim by Engebretson and Moyer (2003) that TRFLP is a robust technique for identifying dominant populations and for calculating diversity statistics in low-diversity communities. Furthermore, these findings verify that TRFLP will miss the majority of low-abundance taxa in highly diverse communities.

## Experimental procedures

### Sample collection

As described previously (Santelli *et al.*, 2008) the basaltic lava samples for this study were collected from basalt outcrops around the big island of Hawai'i during cruise KOK 02-24 on R/V *Ka'imikai-o-Kanaloa* in November 2002 using the manned *Pisces V* submersible. Rock samples were placed in bioboxes containing distilled freshwater, allowed to fill with ambient bottom seawater, sealed to minimize contamination from the upper water column, and brought to the surface for immediate processing. Upon retrieval shipboard, basalt samples were

handled aseptically using flame- and/or ethanol-sterilized equipment. One portion of the sample was stored at  $-80^{\circ}\text{C}$  in a sterile Whirlpak bag for 16S rRNA clone library construction at the Woods Hole Oceanographic Institute; the other portion was stored in sterile cryovials at  $-80^{\circ}\text{C}$  for TRFLP analysis at the Scripps Institution of Oceanography.

### TRFLP analysis

DNA was extracted from samples for TRFLP analysis using the QBioGene FASTDNA SPIN kit for Soil (<http://www.QBioGene.com>, Catalogue #6560-200) and purifying/desalting resultant DNA with a microcon centrifugal filter device (<http://www.millipore.com>, Cat# 42416). The 16S rRNA gene was amplified via polymerase chain reaction (PCR) using universal PCR primers with an annealing temperature of  $56^{\circ}\text{C}$ . The forward primer 68F (5-TNA NAC ATG CAA GTC GRR CG) was fluorescently labelled on the 5' end with 6-FAM (6-carboxyfluorescein); the reverse primer was 1492R (5-RGY TAC CTT GTT ACG ACT T). PCR amplification products were visualized and assayed for size by 1% gel electrophoresis against a 1 kb ladder DNA standard. Fluorescently labelled PCR products were then digested for 6–8 h with various restriction enzymes: AluI, HaeIII, HhaI, MboI, MspI and RsaI (New England Biolabs, Beverly, Mass.). The end-labelled SSU rDNA fragments were separated by polyacrylamide gel electrophoresis against the Genescan-500 ROX size standard with an ABI model 377 automated DNA sequencer, and the data were analysed with the Genescan software (Applied Biosystems, Foster City, Calif.). Electrophoretic resolution of the TRFLP fragments ranged from 50 to 500 bp in length. For the application of richness estimators, it was assumed that one PCR amplicon corresponded to the threshold detection limits of the TRFLP analysis; all peaks were then divided by this value to calculate the frequency or abundance of the OTUs. Two threshold detection limits were evaluated: a peak height of 15 fluorescence units, which corresponds most closely to the signal-to-noise ratio observed in the TRFLP analyses in this study, as well as a higher cut-off criteria of 50 fluorescence units, which is more commonly used in the literature (e.g. Danovaro *et al.*, 2006; Culman *et al.*, 2008).

### 16S rRNA gene clone library construction and *in silico* TRFLP analysis

Procedures for creating 16S rRNA gene clone libraries from the two samples have been presented previously (Santelli *et al.*, 2008). Briefly, DNA was extracted from the samples using the Ultraclean soil DNA kit (MoBio Laboratories) following a slightly modified manufacturer's protocol. Prior to bead beating, the mixture was incubated for 10 min at  $70^{\circ}\text{C}$  after which 200  $\mu\text{g}$  of polyadenylic acid (poly A) was added. The 16S rRNA gene was amplified using the bacteria-specific primer 8F (5'-AGA GTT TGA TCC TGG CTC AG-3') and universal primer 1492R (5'-GGT TAC CTT GTT ACG ACT T-3'). The sequences used in this study (from Santelli *et al.*, 2008) are deposited in GenBank under the Accession Numbers EU491020-EU491090 (from Lō'ihi South Rift) and EU491091-EU491336 (from Lō'ihi Pisces Peak). The com-

puter program DOTUR (Schloss and Handelsman, 2005) was used to calculate species richness estimators from genetic sequence distances using a 97% sequence similarity definition. Using the add-on program tRF-cut (Ricke *et al.*, 2005) for ARB (Ludwig *et al.*, 2004), an *in silico* TRFLP digestion of the nearly full-length 16S rRNA sequences was performed using the restriction enzyme pre-sets in the program for *AluI*, *HaeIII*, *HhaI*, *MboI*, *MspI* and *RsaI* assuming that the 68F primer as above was the labelled primer. The predicted fragments were then grouped by base-pair length to generate abundance curves replicating a TRFLP pattern. Fragments of the same size were grouped into one OTU; the frequency of each OTU was assumed from the number of sequences in each OTU.

#### Calculation of diversity indices

The number and abundance of OTUs calculated in the various data sets were used to evaluate classical non-parametric richness and evenness indices, including the ACE estimator (Chao and Lee, 1992), the Chao1 estimator (Chao, 1984), the Shannon–Weaver index of diversity ( $H'$ ), the Shannon evenness index ( $J'$ ) and the Simpson index of diversity ( $D$ ).  $H'$ ,  $J'$  and  $D$  were calculated as formulated in Schloss and Handelsman (2005).

#### Acknowledgements

We thank D. Rogers and E. Banning for shipboard and laboratory sample processing, R. Davis for comments on the manuscript, C. Moyer for assistance in TRFLP instruction, M. Sogin for sequence analysis pipeline support, *Pisces* pilot T. Kerby for submarine operations and the captain and crew of the *R/V K'O'K*. This work was supported by funding from the NSF Microbial Observatories program grants to K.J.E. (MCB-0348425) and B.M.T. and H.S. (MCB-0348668), by NSF OCE-0526285 and a NASA CAN award to K.J.E., by NSF OCE-0433692 and a NOAA-NURP award to H.S. and B.M.T., and by support from the Agouron Institute to H.S. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these agencies.

#### References

- Bent, S.J., Pierson, J.D., and Forney, L.J. (2007) Measuring species richness based on microbial community fingerprints: the Emperor has no clothes. *Appl Environ Microbiol* **73**: 2399–2401.
- Biddle, J.F., Lipp, J.S., Lever, M., Lloyd, K.G., Sørensen, K.B., Anderson, K., *et al.* (2006) Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Pro Natl Acad Sci U S A* **103**: 3846–3851.
- Blackwood, C.B., Marsh, T., Kim, S.H., and Paul, E.A. (2003) Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. *Appl Environ Microbiol* **69**: 929–932.
- Blackwood, C.B., Hudleston, D., Zak, D.R., and Buyer, J.S. (2007) Interpreting ecological diversity indices applied to terminal restriction fragment length polymorphism data: insights from simulated microbial communities. *Appl Environ Microbiol* **73**: 5276–5283.
- Chao, A. (1984) Non-parametric estimation of the number of classes in a population. *Scand J Stat* **11**: 265–270.
- Chao, A., and Lee, S.M. (1992) Estimating the number of classes via sample coverage. *J Am Stat Assoc* **87**: 210–217.
- Clement, B.G., Kehl, L.E., Debord, K.L., and Kitts, C.L. (1998) Terminal restriction fragment patterns (TRFPs), a rapid, PCR-based method for the comparison of complex bacterial communities. *J Microbiol Methods* **31**: 135–142.
- Colwell, R.K., and Coddington, J.A. (1994) Estimating terrestrial biodiversity through extrapolation. *Phil Trans R Soc Lond B* **345**: 101–118.
- Culman, S.W., Gauch, H.G., Blackwood, C.B., and Thies, J.E. (2008) Analysis of T-RFLP data using analysis of variance and ordination methods: a comparative study. *J Microbiol Methods* **75**: 55–63.
- Danovaro, R., Luna, G.M., Dell'Anno, A., and Pietrangeli, B. (2006) Comparison of two fingerprinting techniques, terminal restriction fragment length polymorphism and automated ribosomal intergenic space analysis, for determination of bacterial diversity in aquatic environments. *Appl Environ Microbiol* **72**: 5982–5989.
- Dunbar, J., Takala, S., Barns, S.M., Davis, J.A., and Kuske, C.R. (1999) Levels of bacterial community diversity in four arid soils compared by cultivation and 16S rRNA gene cloning. *Appl Environ Microbiol* **65**: 1662–1669.
- Dunbar, J., Ticknor, L.O., and Kuske, C.R. (2000) Assessment of microbial diversity in four southwestern United States soils by 16S rRNA gene terminal restriction fragment analysis. *Appl Environ Microbiol* **68**: 2943–2950.
- Engebretson, J.L., and Moyer, C.L. (2003) Fidelity of select restriction endonucleases in determining microbial diversity by terminal-restriction fragment length polymorphism. *Appl Environ Microbiol* **68**: 4823–4829.
- Hartmann, M., Frey, B., Kölliker, R., and Widmer, F. (2005) Semi-automated genetic analyses of soil microbial communities: comparison of T-RFLP and RISA based on descriptive and discriminative statistical approaches. *J Microbiol Methods* **61**: 349–360.
- Huber, J.A., Mark Welch, D.B., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., and Sogin, M.L. (2007) Microbial population structures in the deep marine biosphere. *Science* **318**: 97–100.
- Hughes, J.B., Hellmann, J.J., Ricketts, T.H., and Bohannan, B.J.M. (2001) Counting the uncountable: statistical approaches to estimating microbial diversity. *Appl Environ Microbiol* **67**: 4399–4406.
- Hughes Martiny, J.B., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., *et al.* (2006) Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**: 102–112.
- Inagaki, F., Nunoura, T., Nakagawa, T., Teske, A., Lever, M., Lauer, A., *et al.* (2006) Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Pro Natl Acad Sci U S A* **103**: 2815–2820.
- Kaplan, C.W., and Kitts, C.L. (2003) Variation between observed and true terminal restriction fragment length is

- dependent on true TRF length and purine content. *J Microbiol Methods* **54**: 121–125.
- Lipp, J.S., Morono, Y., Inagaki, F., and Hinrichs, K.-U. (2008) Significant contribution of *Archaea* to extant biomass in marine subsurface sediments. *Nature* **454**: 991–994.
- Liu, W.T., Marsh, T.L., Cheng, H., and Forney, L.J. (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol* **63**: 4516–4522.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, *et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- Ricke, P., Kolb, S., and Braker, G. (2005) Application of a newly developed ARB software-integrated tool for in silico terminal restriction fragment length polymorphism analysis reveals the dominance of a novel pmoA cluster in forest soil. *Appl Environ Microbiol* **61**: 1671–1673.
- Santelli, C., Orcutt, B., Banning, E., Moyer, C., Bach, W., Staudigel, H., and Edwards, K. (2008) Abundance and diversity of microbial life in the ocean crust. *Nature* **453**: 653–656.
- Schippers, A., and Neretin, L.N. (2006) Quantification of microbial communities in near-surface and deeply buried marine sediments on the Peru continental margin using real-time PCR. *Environ Microbiol* **8**: 1251–1260.
- Schloss, P., and Handelsman, J. (2005) Introducing DOTUR: a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501–1506.
- Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., *et al.* (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci U S A* **103**: 12115–12120.
- Sørensen, K.B., and Teske, A. (2006) Stratified communities of active archaea in deep marine subsurface sediments. *Appl Environ Microbiol* **72**: 4596–4603.
- Youssef, N.H., and Elshahed, M.S. (2008) Species richness in soil bacterial communities: a proposed approach to overcome sample size bias. *J Microbiol Methods* **75**: 86–91.