INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
MICROBIAL DIVERSITY AND COMMUNITY STRUCTURE
DETERMINATIONS THROUGH ANALYSES OF SSU rRNA GENE
DISTRIBUTIONS AND PHYLOGENY

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

OCEANOGRAPHY

DECEMBER 1995

By

Craig Lee Moyer

Dissertation Committee:

David M. Karl, Chairperson
Fred C. Dobbs
Michael R. Landry
Robert R. Bidigare
Neil S. Reimer
ACKNOWLEDGMENTS

I would first like to express my sincere gratitude for the wisdom, patience, and knowledge passed along to me by my mentor David Karl. I have learned much by the example of excellence he consistently sets, and only hope that some of it will stick with me for the duration. I am also grateful to him for showing me what it was like to dive to the ocean's depths inside a submersible and its potential as a platform for scientific research. Likewise, I shall always remain grateful to my former mentor Richard Morita, who also bore the burden of my tenure in his laboratory. In addition, I acknowledge Fred Dobbs, whom has acted as a co-mentor during my Ph.D. program. He has spent much time and effort in the capacity as editor and helped immensely with much of my field work at Loihi. Without all of their continual inspiration, help, and advice I would never have discovered some of the wonders that planet ocean has to offer, let alone achieved this degree.

This dissertation project would not have been possible without the support and contributions of the rest of my Ph.D. committee members, who were integral in aiding with the direction of my research and the shaping of this document. I wish to especially acknowledge the helpful assistance, technical expertise, and thoughtful discussions of Neil Reimer and Lisa Sampson of the University of Hawaii Biotechnology-Molecular Biology Instrument and Training Facility. Special thanks are due to the captain and crew of the R/V Kila, the DSRV Pisces V operations team, and the staff of the Hawaiian Undersea Research Laboratory. I also gratefully acknowledge Edward DeLong for his technical assistance and encouragement with the phylogenetic sequence data analysis.
This project was funded by a grant from the National Oceanic and Atmospheric Administration, project R/OM-8, which is sponsored by the University of Hawaii Sea Grant College Program (SOEST), under Institutional Grant NA89AA-D-SG063 from the National Oceanic and Atmospheric Administration Office of Sea Grants; by the National Oceanic and Atmospheric Administration-National Undersea Research Program, Department of Commerce; and by a Research and Training Revolving Fund Award from the University of Hawaii Research Council.
ABSTRACT

The majority of this study focused on the analysis of diversity and community structure through an examination of small-subunit (SSU) rRNA genes from the microbial mat assemblage located at an active, hydrothermal vent system, Loihi Seamount, Hawaii. The habitat studied was Pele’s Vents, a deep-sea hydrothermal ecosystem located near the summit of Loihi. Through a restriction fragment length polymorphism (RFLP) distribution analysis of a SSU rDNA clone library generated by the polymerase chain reaction (PCR), it was determined that two operational taxonomic units (OTUs) dominated the bacterial assemblage of the mat community. These two OTUs together accounted for ~73% of the bacterial SSU rDNA clone library examined, with ten OTUs accounting for the remaining ~27%. A technique, analogous to rarefaction, was developed to determine that diversity had been sufficiently described by the clones examined. SSU rDNA fingerprinting of clones belonging to each OTU was conducted to confirm OTU specificity and SSU rDNA identity.

The phylogenetic diversity of the bacterial OTUs discovered at Pele’s Vents (abbreviated PVB, for Pele’s Vents Bacteria) was also described. Two dominant phylotypes were found that included more than a single OTU, and therefore comprised a cluster of phylogenetically related taxa. The most abundant phylotype was the PVB OTU 2 cluster, which was comprised of PVB OTUs 2, 3, 6, and 8. The PVB OTU 2 cluster accounted for ~61% of the bacterial clone library and had a lineage contained in the \(\varepsilon\)-Proteobacteria subclass and a Thiovulum-like phylogeny. The second dominant
The phylogenetic diversity of the archaeal OTUs discovered at Pele's Vents (abbreviated PVA, for Pele's Vents Archaea) were also described. The archaean clones were dominated by the PVA OTU 2 cluster (comprised of PVA OTUs 2, 3, and 4), which was phylogenetically contained in the Crenarchaeota, whereas PVA OTU 1 was contained in the Euryarchaeota. Both of these lineages were phylogenetically affiliated with recently discovered cosmopolitan marine archaeoplankton.
# TABLE OF CONTENTS

**ACKNOWLEDGMENTS** ........................................................................................................ iv

**ABSTRACT** ...................................................................................................................... vi

**LIST OF TABLES** ........................................................................................................... x

**LIST OF FIGURES** ........................................................................................................ xi

**CHAPTER 1: INTRODUCTION AND RATIONALE FOR THE PRESENT RESEARCH** .......................................................... 1  
  **BACKGROUND** ........................................................................................................ 1  
  **SYNOPSIS** .................................................................................................................. 11  
  **REFERENCES** ............................................................................................................ 14

**CHAPTER 2: A COMPUTER-SIMULATED RFLP ANALYSIS OF BACTERIAL SSU rRNA GENES: EFFICACY OF SELECTED TETRAMERIC RESTRICTION ENZYMES** 20  
  **ABSTRACT** ............................................................................................................. 20  
  **INTRODUCTION** ...................................................................................................... 21  
  **MATERIALS AND METHODS** .................................................................................. 23  
    - Selection of bacterial sequence data included in model ........................................ 23  
    - Computer generation of RFLPs ........................................................................... 24  
    - Phylogenetic analyses of RFLP and DNA sequence data .................................. 25  
  **RESULTS** ................................................................................................................... 26  
  **DISCUSSION** .......................................................................................................... 41  
  **REFERENCES** ............................................................................................................ 45

**CHAPTER 3: ESTIMATION OF DIVERSITY AND COMMUNITY STRUCTURE THROUGH RFLP DISTRIBUTION ANALYSIS OF BACTERIAL 16S rRNA GENES FROM A MICROBIAL MAT AT AN ACTIVE, HYDROTHERMAL VENT SYSTEM, LOIHI SEAMOUNT, HAWAII** ........................................................................................................... 47  
  **ABSTRACT** ............................................................................................................. 47  
  **INTRODUCTION** ...................................................................................................... 48  
  **MATERIALS AND METHODS** .................................................................................. 51  
    - Collection of bacterial mat samples ................................................................. 51  
    - Genomic DNA extraction and purification .................................................... 52  
    - Oligonucleotide synthesis and purification .................................................. 54  
    - Amplification of 16S rRNA genes .................................................................. 55  
    - Construction of bacterial 16S rDNA clone library ....................................... 56  
    - 16S rDNA RFLP analysis .................................................................................. 57  
    - 16S rDNA fingerprinting by Southern blotting and oligonucleotide probe hybridization ............................................................... 58
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Summary of Loihi hydrothermal fluid chemistry and comparison with other known deep-sea vent habitats</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Average number of restriction sites per taxon for each tetrameric restriction enzyme</td>
<td>34</td>
</tr>
<tr>
<td>2.2</td>
<td>Number of phylogenetically estimated bootstrap values for each tetrameric restriction enzyme (TRE) group</td>
<td>36</td>
</tr>
<tr>
<td>2.3</td>
<td>Percent (%) “successful” phylogenetic affiliations across tetrameric restriction enzyme (TRE) groups as determined by phylogenetic analysis of RFLP data</td>
<td>40</td>
</tr>
<tr>
<td>3.1</td>
<td>Data matrix for OTU polymorphisms as detected by restriction digest</td>
<td>61</td>
</tr>
<tr>
<td>4.1</td>
<td>Phylogenetic affiliations and percent recoveries of bacterial SSU rRNA genes from a microbial mat at Pele’s Vents</td>
<td>86</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Universal phylogenetic tree in rooted form, showing the three domains, <em>Archaea</em>, <em>Bacteria</em>, and <em>Eucarya</em></td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>Restriction fragment-size frequency distributions cumulative across all taxa for the tetrameric restriction enzymes <em>AluI</em>, <em>DdeI</em>, and <em>MspI</em> (Group 1)</td>
<td>28</td>
</tr>
<tr>
<td>2.2</td>
<td>Restriction fragment-size frequency distributions cumulative across all taxa for the tetrameric restriction enzymes <em>MboI</em>, <em>HinfI</em>, <em>TaqI</em>, and <em>HaeIII</em> (Group 2)</td>
<td>30</td>
</tr>
<tr>
<td>2.3</td>
<td>Restriction fragment-size frequency distributions cumulative across all taxa for the tetrameric restriction enzymes <em>HhaI</em>, <em>RsaI</em>, and <em>BstUI</em> (Group 3)</td>
<td>32</td>
</tr>
<tr>
<td>2.4</td>
<td>Phylogenetic tree based on the SSU rRNA sequence data for all bacterial taxa examined (see Appendix B) using the neighbor-joining distance matrix method</td>
<td>38</td>
</tr>
<tr>
<td>3.1</td>
<td>Map showing the location of Loihi Seamount in relation to the island of Hawaii</td>
<td>53</td>
</tr>
<tr>
<td>3.2</td>
<td>OTUs determined from RFLP analysis of representative cloned inserts of bacterial 16S rDNA restricted with <em>AluI</em> plus <em>RsaI</em></td>
<td>62</td>
</tr>
<tr>
<td>3.3</td>
<td>16S rDNA fingerprinting of RFLPs from clones representing OTUs with oligonucleotide probe 536F</td>
<td>63</td>
</tr>
<tr>
<td>3.4</td>
<td>Distribution in OTUs of 48 bacterial 16S rDNA clones from the bacterial mat community at an active, hydrothermal vent system, Loihi Seamount, Hawaii</td>
<td>65</td>
</tr>
<tr>
<td>3.5</td>
<td>Estimation of diversity in the bacterial mat community at an active, hydrothermal vent system, Loihi Seamount, Hawaii</td>
<td>66</td>
</tr>
<tr>
<td>4.1</td>
<td>Phylogenetic tree demonstrating relationships of the PVB OTU 2 cluster with <em>ε-Proteobacteria</em> as determined by maximum likelihood analysis of SSU rDNA sequences</td>
<td>89</td>
</tr>
</tbody>
</table>
4.2 Phylogenetic tree demonstrating relationships of the PVB OTU 1 cluster, PVB OTU 5, and PVB OTU 12 with \( \gamma \)-Proteobacteria as determined by maximum likelihood analysis of SSU rDNA sequences ................................. 92

4.3 Phylogenetic tree demonstrating relationships of PVB OTU 4 with \( \gamma \)-Proteobacteria as determined by maximum likelihood analysis of SSU rDNA sequences ................................. 95

4.4 Phylogenetic tree demonstrating relationships of PVB OTU 10B with other myxobacteria as determined by maximum likelihood analysis of SSU rDNA sequences ................................. 97

4.5 Phylogenetic tree demonstrating relationships of PVB OTU 9A with the Paraphyletic assemblage and the Leptospirillum group as determined by maximum likelihood analysis of SSU rDNA sequences ......................................................... 99

5.1 Phylogenetic tree demonstrating relationships the of the PVA OTU 2 cluster with the marine archaeal Group I and PVA OTU 1 with the marine archaeal Group II as determined by the DeSoete distance matrix algorithm ......................... 121
CHAPTER 1

INTRODUCTION AND RATIONALE FOR THE PRESENT RESEARCH

BACKGROUND

One of the long-standing challenges in microbial ecology is the identification and description of microbial populations within their communities without the technical limitations imposed by cultivation methods. These understandings are critical for determining the vast levels of environmental microbial diversity, which remains one of our greatest untapped natural resources. To address this challenge, recently developed molecular biological techniques were applied herein to ascertain the structure and diversity of microbial communities. The overall endeavor was to examine specific habitats for microorganisms, and consequently sites of biogeochemical activity. The primary focus of the studies conducted herein was on the community that comprised the microbial mats from the deep-sea hydrothermal vent system known as Pele’s Vents, located at the summit of Loihi Seamount, Hawaii.

The discovery of deep-sea hydrothermal vent systems and the unexpected presence of dense, but spatially-restricted microbial and animal communities within the fields of hydrothermal fluid discharge occurred first at the Galapagos Rift zone (Corliss et al., 1979; Grassle et al., 1979; Lonsdale, 1977). These initial submersible-based discoveries have been followed by nearly two decades of exploration and experimentation. Numerous
additional discoveries in the Pacific and Atlantic Ocean basins have clearly established the importance of vent systems as sources of mantle-derived gases and solutes, and as loci for heat dissipation, polymetallic sulfide mineral deposition, geochemical exchange, and microbial activity.

Most of the hydrothermal systems discovered to date have been at, or near, tectonic plate boundaries. However, data derived from dredging (DeCarlo et al., 1983; Malahoff et al., 1982), near-bottom water-sampling (Gamo et al., 1987; Horibe et al., 1983; Sakai et al., 1987), seafloor camera survey (Malahoff et al., 1987), and submersible dive programs (Karl et al., 1988a,b) have confirmed that hydrothermal vents are present at the summit of Loihi Seamount, a mid-plate, hotspot submarine volcano located at the southern terminus of the Hawaiian-Emperor Island chain. Loihi Seamount is the youngest feature of this island chain, which spans a distance of about 2,600 km and has an age of about 70 million years. Loihi Seamount is located 35 km southeast of the Big Island of Hawaii, and rises ~4,000 m above the seafloor to within ~1,000 m of the sea surface (Fig. 3.1).

During the initial DSRV ALVIN and PISCES V submersible dives on Loihi, a hydrothermal vent field was discovered on the upper portion of the south rift zone near the summit of the volcano (Karl et al., 1988a,b). This vent field (named "Pele's Vents" after the Hawaiian volcano goddess) has an area of approximately 250 m², with individual vents that span a few to twenty centimeters across producing discharge velocities of 1-10 cm sec⁻¹. The lava from the venting area is brecciated, strongly vesicular (20-35 vol%), weakly phryic (1-2 vol% plagioclase, olivine, and augite), alkalic in composition, and degases vigorously when brought to the surface. The lava in and around
the vents is mantled with red-orange "nontronite" deposits, an assemblage of varying amounts of iron oxides, silica, and Fe-rich smectite (DeCarlo et al., 1983; Malahoff et al., 1982). Thick, orange-stained microbial mats with iron oxide coatings cover the lava on the periphery of the vent field (Karl et al., 1988a,b). White "streamers" of bacteria, together with white deposits of elemental sulfur are present on the lava surfaces adjacent to active vents.

The number of Pacific seamounts originating from hot-spots has been conservatively estimated at 1.5 × 10⁶ (Fornari et al., 1987). This estimate continues to rise with the advent of wider swath coverage using multibeam sonar systems. Hot spot volcanism has been recognized as a major source of magma to the oceanic crust, as it generates about 10% of the new oceanic crust in the form of alkaline and tholeiitic basalts and rhyolite. Recent intrusion of magma into Loihi Seamount is suggested by the presence of fresh lavas, seismic activity, and emissions of hydrothermal fluids from diffuse low-temperature (<35°C) vents (Appendix A; Edmond et al., 1987; Karl et al., 1988a,b; Karl et al., 1989; Malahoff et al., 1982; Malahoff et al., 1987; Sedwick et al., 1992). These attributes contribute to the chemistry of the vent waters from Pele's Vents, which has both similarities to and differences from other low-temperature vent fields, such as those found on the Galapagos Rift and on Axial Seamount, a hotspot volcano located on the Juan de Fuca Ridge (Table 1.1). Dissolved SiO₂ positively correlates with vent temperature (Sedwick et al., 1992; Von Damm et al., 1985), suggesting that silica can be used as a conservative tracer for mixing of the vent fluids with ambient seawater. Other vent water components also display linear trends with SiO₂, suggesting a common source for the
Table 1.1. Summary of Loihi hydrothermal vent fluid chemistry and comparison with other known deep-sea vent habitats (From Karl et al., 1988b).

<table>
<thead>
<tr>
<th>Component</th>
<th>Pele's Vents</th>
<th>Ambient seawater</th>
<th>Venti water enrichment (+) or depletion (-) (%)</th>
<th>Axial Seamoun</th>
<th>Galapagos Rift composite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li (μM)</td>
<td>36.5</td>
<td>23.6</td>
<td>+55</td>
<td>58.0</td>
<td>79.7-116.6</td>
</tr>
<tr>
<td>Na (mM)</td>
<td>464</td>
<td>455</td>
<td>+2</td>
<td>—</td>
<td>10.7</td>
</tr>
<tr>
<td>K (mM)</td>
<td>12.3</td>
<td>10.2</td>
<td>+21</td>
<td>—</td>
<td>1.28-2.89</td>
</tr>
<tr>
<td>Rb (μM)</td>
<td>2.85</td>
<td>1.25</td>
<td>+128</td>
<td>4.4</td>
<td>11.5-12.7</td>
</tr>
<tr>
<td>Mg (mM)</td>
<td>50.3</td>
<td>52.2</td>
<td>—</td>
<td>49.39</td>
<td>48.4</td>
</tr>
<tr>
<td>Ca (mM)</td>
<td>11.2</td>
<td>10.3</td>
<td>+9</td>
<td>11.06</td>
<td>11.5-12.7</td>
</tr>
<tr>
<td>Sr (μM)</td>
<td>88.5</td>
<td>86.9</td>
<td>+2</td>
<td>92</td>
<td>86.6-87.3</td>
</tr>
<tr>
<td>Ba (μM)</td>
<td>0.71</td>
<td>0.07</td>
<td>+14</td>
<td>1.21</td>
<td>1.52-1.53</td>
</tr>
<tr>
<td>Si (mM)</td>
<td>1,285</td>
<td>123</td>
<td>+945</td>
<td>1,100</td>
<td>1,884</td>
</tr>
<tr>
<td>Cl (mM)</td>
<td>524</td>
<td>534</td>
<td>—</td>
<td>531.9</td>
<td>524-545</td>
</tr>
<tr>
<td>SO₂ (mM)</td>
<td>27.0</td>
<td>28.6</td>
<td>—</td>
<td>25.94</td>
<td>26.2</td>
</tr>
<tr>
<td>Fe (μM)</td>
<td>1,010</td>
<td>0.05</td>
<td>+2 x 10⁶</td>
<td>2.6</td>
<td>245</td>
</tr>
<tr>
<td>Mn (μM)</td>
<td>21.1</td>
<td>0.01</td>
<td>+2 x 10⁵</td>
<td>27.7</td>
<td>36.2-115</td>
</tr>
<tr>
<td>NH₄ (μM)</td>
<td>7.254#</td>
<td>7.2</td>
<td>+1 x 10⁵</td>
<td>—</td>
<td>6,910-24,200**</td>
</tr>
<tr>
<td>NO₃ (μM)</td>
<td>5.38</td>
<td>0.72</td>
<td>+619</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>NO₂ (μM)</td>
<td>3.46</td>
<td>2.93</td>
<td>+18</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>[NO₂⁺ + NO₃⁻] (μM)</td>
<td>11.9</td>
<td>42.1</td>
<td>—</td>
<td>0.0**</td>
<td>—</td>
</tr>
<tr>
<td>CO₃²⁻ (mM)</td>
<td>300†</td>
<td>2.38</td>
<td>+1.3 x 10⁵</td>
<td>40</td>
<td>2.60-3.54†</td>
</tr>
<tr>
<td>pH</td>
<td>5.3-5.55‡</td>
<td>7.51</td>
<td>+1.6 x 10⁴</td>
<td>6.18</td>
<td>6.65</td>
</tr>
</tbody>
</table>

* Water samples for major and minor inorganic constituents were processed on board ship by filtration through prewashed 0.2 μm Nuclepore filters followed by pH adjustment to pH 2.0 with ultrapure nitric acid. Na, K, Rb, Ca, Mg, Si, and for concentrated vent waters, Fe and Mn were determined by flame atomic absorption spectrophotometry. Minimum precisions based on multiple determinations of both samples and standards are: Na (1%), K (1%); Rb (10%); Ca (2%); Mg (2%); Si (3%); Fe (2%); Mn (2%). Lithium (4%) was determined by flame atomic emission spectrophotometry. Ba, Fe and Mn in ambient waters were determined by graphite furnace atomic absorption spectrophotometry, following preconcentration for Fe and Mn. Minimum precisions are: Ba (30%), Fe (10%), Mn (10%). SO₂ (1.5%) was determined by ion chromatography. Chloride (0.1%) was determined by potentiometric titration with AgNO₃. For dissolved nutrient determinations, water samples were filtered through combusted glass fibre filters and stored frozen for subsequent analysis by standard autoanalyser procedures.

† Sample collected with a 5-litre Niskin bottle at 980 m near Pele’s Vents; water temperature = 3.5°C.

§ Extrapolated values to 30°C from T versus Si data of Edmond et al.23.
| Sr analyses by isotope dilution, mass spectrometry courtesy of J. D. Macdougall.
| Highest value at 9°C for Clamabake Field; all other Galapagos Rift fields display mid-range Fe maxima and decrease at higher temperatures.

Reported as total dissolved metal. From Edmond et al.33.

# Extrapolated value to 1,285 μM Si from CH₄ versus Si data of Karl et al.44.

** Extrapolated values to 30°C from data of Lilley et al.20.

†† Extrapolated value to 1,285 μM Si from CO₂ versus Si data of Karl et al.14.

†‡ Extrapolated values to 30°C from data of Edmond et al.24.

§§ From Edmond et al.33.

[1] From Gamo et al.27.
numerous individual vents emanating from the Pele's vent field. Relative to deep seawater, the vent waters are enriched in total dissolved CO$_2$, CH$_4$, NH$_4^+$, PO$_4^{3-}$, Fe, Mn, Li, K, Rb, Ca, Sr, and Ba and depleted in Mg, SO$_4^{2-}$, [NO$_3^-$+NO$_2^-$], and Cl$^-$ (Table 1.1; Karl et al., 1988a,b; Karl et al., 1989). The most notable of the distinctive chemical characteristics of these hydrothermal vent waters are the presence of high levels of total dissolved CO$_2$ (>300 mM) and dissolved iron (~1 mM), and the relatively low level of measurable dissolved H$_2$S (~2 μM), with a resulting \textit{in situ} pH as low as 4.2 (Edmond et al., 1987; Karl et al., 1988a,b; Karl et al., 1989; Sedwick et al., 1992).

It has been hypothesized that the internal heat of our planet, maintained through the radioactive decay of long-lived isotopes of uranium, thorium, and potassium, may represent an alternative energy source to incoming solar radiation, e.g., via chemosynthesis rather than photosynthesis (Jannasch and Mottl, 1985). A rigorous experimental evaluation of the existence and quantitative significance of these fundamentally different sources of energy for microbial nutrition is an objective of the highest priority (Karl, 1987; Karl, 1995). Hydrothermal vent water samples (15-30°C) collected at Pele's Vents contained elevated concentrations of particulate ATP (up to 118 ng l$^{-1}$) relative to ambient bottom seawater (Karl et al., 1989). These ATP values are comparable to the enrichments measured at other deep-sea hydrothermal vent sites (Karl, 1987; Karl, 1995; Karl et al., 1980), and indicate the presence of a localized source of microorganisms. Microbes present in the water samples collected from the warm vent waters actively oxidized $^{14}$CH$_4$ to $^{14}$CO$_2$ and incorporated $^3$H-adenine and $^3$H-thymidine into nucleic acids when incubated at 30°C, and 1 atm, indicating active metabolism and cell growth (Karl et al., 1988a; Karl et al., 1989).
The presence of active thermophilic bacteria with high rates of metabolic activity at 60°C, and 1 atm, was also detected (Karl et al., 1989). Based on these initial observations and sample collections, there appeared to be at least four distinct microbial habitats at Loihi Seamount: (i) the anoxic subsurface vent system, (ii) the region immediately surrounding the vent loci where the anoxic hydrothermal vent fluids are mixed with oxygenated deep-sea waters, (iii) the region of extensive microbial mat formation that is restricted to the periphery of the main hydrothermal vent field, and (iv) the hydrothermal plume.

Iron and sulfur oxidation are common metabolic processes in the microbial world (e.g., Sulfolobus sp., Leptospirillum sp., Acidiphilium sp., Thiobacillus sp., Sphaerotilus sp., Gallionella sp., Leptothrix sp., Thiomicrospira sp., Thiothrix sp., Beggiatoa sp., Thiovulum sp.\(^1\), Thioploca sp.\(^1\) [Ehrlich, 1981; Ghiorse, 1984; Hanert, 1991; Kuenen et al., 1991; La Rivière and Schmidt, 1991; Nelson, 1991; Robertson and Kuenen, 1991]). Not all sulfur- and iron-oxidizing bacteria couple the oxidation to CO\(_2\) reduction and in some cases the oxidation of these metals is only an indirect metabolic process (i.e., non-enzymatic). At Pele’s Vents, the relatively high concentrations of dissolved ferrous iron in the discharge waters (relative to H\(_2\)S) and low pH should favor the growth of iron-oxidizing bacteria.

Selective enrichment culture has severe limitations as an approach to the cultivation of naturally-occurring microorganisms (Poindexter and Leadbetter, 1986; Ward et al., 1992). The majority (typically >90-99%) of bacteria in nature cannot be cultivated using

---

\(^1\)Note: No pure culture isolates as yet exist for these genera, possibly due to their being obligate chemoaerobicolitrophs with essential syntrophic partners.
traditional techniques (Brock, 1987; Jannasch and Jones, 1959). Consequently, it is very unlikely that collections of bacterial isolates are representative of in situ diversity and community structure. Furthermore, because relatively nutrient-rich media are generally used for isolations, copiotrophic bacteria may be selected for rather than those dominant in the natural community. Obtaining a better understanding of bacterial community structure and diversity is crucial to aspects of microbial ecology where bacteria interact with one another and with their environment, e.g., global biogeochemical cycling of matter, risk assessment related to the release of genetically engineered microorganisms, predator-prey relationships, and trophic-level interactions.

Cell component analyses provide a culture-independent means of investigating microorganisms as they occur in nature (Olsen et al., 1986; Pace et al., 1986; Ward et al., 1992). While several types of cell components have been analyzed, the small-subunit ribosomal RNA (SSU rRNA) molecule offers an amount and type of information that makes it one of the best culture-independent descriptors or biomarkers of microorganisms. For example, each SSU rRNA gene contains highly conserved regions found among all living organisms as well as diagnostic variable regions unique to particular organisms or closely related groups. Additionally, each SSU rRNA gene contains about 1,500 nucleotides of sequence information that can be obtained and utilized to differentiate among closely-related and distantly-related groups of microorganisms. Therefore, this information has phylogenetic meaning and can be analyzed in the context of a large and growing database of SSU rRNA sequences of cultivated and uncultivated microorganisms. The sequence information also provides a systematic approach for developing oligonucleotides
that can be used as hybridization probes to the nucleic acids from microorganisms contributing specific SSU rRNA sequences to natural communities. This type of molecular approach allows the autecology of microorganisms to be studied whether or not they have ever been cultivated (Olsen et al., 1986; Pace et al., 1986; Ward et al., 1992).

In recent years a detailed theory of evolutionary relationships among the domains Bacteria, Archaea and Eucarya has emerged from comparisons of molecular "signature" sequences (Fig. 1.1; Fox et al., 1980; Olsen et al., 1994; Winker and Woese, 1991; Woese, 1987; Woese, 1994; Woese et al., 1990). Ribosomal RNAs have been the most widely used biomolecules for phylogenetic comparisons for the following reasons: (i) they are essential components of the protein synthesis machinery and therefore, are ubiquitously distributed and functionally conserved in all organisms, (ii) they lack the interspecies horizontal gene transfer found with many prokaryotic genes, (iii) they are readily isolated and identified, and (iv) they contain variable regions interspersed among highly conserved regions of primary and secondary structure, permitting phylogenies to be inferred over a broad range of evolutionary distance. These features make rRNAs particularly useful for studies of microbial community ecology, where a broad and unknown diversity of microorganisms is likely to exist. Currently, over 3,000 SSU rRNA sequences have been made available for study from the Ribosomal Database Project (Larsen et al., 1993; Maidak et al., 1994; Olsen et al., 1991; Olsen et al., 1992), which provides these data in an aligned and phylogenetically ordered format.

Relatively few studies have described the composition of naturally-occurring microbial communities. To date only two other environments (i.e., in addition to
Figure 1.1. Universal phylogenetic tree in rooted form, showing the three domains, *Archaea*, *Bacteria*, and *Eucarya*; adapted from Woese (1994). The position of the root was determined by using the paralogous gene couple, translation elongation factors EFTu and EFG (Iwabe *et al.*, 1989).
hydrothermal vents) have been examined extensively, and both have shown high levels of
diversity with respect to uncultured microorganisms. The first environment investigated
was the near surface oligotrophic ocean of the North Atlantic gyre (Britschgi and
Giovannoni, 1991; Giovannoni et al., 1990; Mullins et al., 1995) and the North Pacific
gyre (DeLong et al., 1993; Schmidt et al., 1991), and from both of these gyres at the
surface to depths up to 500 m (Fuhrman et al., 1993). Recently, two novel lineages of
pelagic marine Archaea have been located in coastal picoplankton populations from the
surface to depths up to 500 m (DeLong, 1992; DeLong et al., 1994; Fuhrman et al., 1992),
and have been quantified to represent approximately 1-2% of the total picoplankton off the
coast of Santa Barbara (DeLong, 1992) and as much as 30% near Arthur Harbor,
Antarctica (DeLong et al., 1994).

The second environment to be examined was in Yellowstone National Park, where
microbial mats proliferate in hot spring communities (e.g., Octopus Spring [Ward et al.,
1990; Weller and Ward, 1989; Weller et al., 1991; Weller et al., 1992; Reysenbach et al.,
1994]). As with marine picoplankton, a broad spectrum of diversity was detected and
spanned the domain Bacteria; all cloned isolates represent as yet uncultured taxa.
Recently, in another Yellowstone spring (Jim’s Black Pool) having high concentrations of
reduced iron and sulfide, a large number of Crenarchaeota clones were detected and shown
to represent high levels of phylogenetic diversity within this archaeal kingdom (Barns et al.,
1994).

This dissertation focuses on the identification and description of the assemblage of
microbial populations (i.e., both from the Bacteria and Archaea domains; sensu Woese
from a microbial mat community within an active hydrothermal vent system, Pele's Vents, located near the summit of the submarine volcano Loihi. The research objectives were achieved by combining the precise sampling capabilities of a deep-sea research submersible with modern molecular biological techniques that take advantage of SSU rRNA genes. These genes were used as discriminators for categorizing multiple microbial populations and as descriptors of a single microbial population's evolutionary history or phylogeny.

SYNOPSIS

This dissertation is divided into six chapters, of which this introduction is the first. Chapter 2 describes a computer-simulated RFLP analysis of bacterial SSU rRNA genes using several tetrameric restriction enzymes in a effort to determine which of these enzymes were most efficient at discriminating among bacterial taxa. Three general distribution patterns were observed for the RFLPs generated by each tetrameric restriction enzyme. The analysis was based on these distribution patterns and the estimated phylogenetic affiliations associated with these groups of restriction enzymes. The restriction enzymes HhaI, Rsal, and BstUI (and their respective isoschizomers) were determined as the most efficacious for detecting and separating bacterial SSU rRNA genes. Chapter 2 represents a manuscript submitted to *Applied and Environmental Microbiology* under the authorship of C. L. Moyer, J. M. Tiedje, F. C. Dobbs, and D. M. Karl.

Chapter 3 constitutes the initial estimation of community structure and diversity.
through the use of a RFLP distribution analysis of bacterial SSU rRNA genes from the
microbial mats at Pele's Vents. It was determined that two OTUs (1 and 2) dominated the
bacterial mat community, accounting for ~73% of the bacterial clone library. The
remaining ten OTUs (which had no more than three clones per OTU) accounted for ~27%
of the bacterial clone library. A technique analogous to rarefaction was developed to
determine whether the bacterial diversity had been sufficiently described by the clones
examined. SSU rDNA fingerprinting of clones belonging to each OTU was conducted to
confirm OTU specificity and SSU rDNA identity. Chapter 3 represents a manuscript that
was published in the journal *Applied and Environmental Microbiology* under the
authorship of C. L. Moyer, F. C. Dobbs, and D. M. Karl (Moyer et al., 1994).

Chapter 4 reports on the phylogenetic diversity of the PVB OTUs (abbreviated
PVB, for Pele's Vents *Bacteria*), which were determined in the previous study. This
analysis demonstrated that phylogenetically similar OTUs were present for each of the two
dominant OTUs (1 and 2) constituting the PVB OTU 1 and 2 clusters. The most abundant
phylotype was the PVB OTU 2 cluster, which was comprised of the PVB OTUs 2, 3, 6,
and 8. The PVB OTU 2 cluster accounted for ~61% of the bacterial clone library and had
a lineage contained in the e-Proteobacteria subclass and a *Thiovolum*-like phylogeny. The
other dominant phylotype was the PVB OTU 1 cluster, which was comprised of
PVB OTUs 1 and 11 and accounted for ~27% of the bacterial clone library. The
PVB OTU 1 cluster had a lineage within the γ-Proteobacteria subclass and a
*Xanthomonas*-like phylogeny. The remaining five PVB OTUs, each determined by a single
clone, were represented by a wide variety of phylotypes spanning the domain *Bacteria.*
Chapter 4 represents a manuscript that was published in the journal *Applied and Environmental Microbiology* under the authorship of C. L. Moyer, F. C. Dobbs, and D. M. Karl (Moyer *et al.*, 1995a).

Chapter 5 is a continuation of the microbial community assessment at Pele's Vents hydrothermal system. The phylogenetic diversity of archaeal OTUs (abbreviated PVA, for Pele's Vents *Archaea*) discovered at this habitat were described. The archaeal SSU rDNA clones were dominated by the PVA OTU 2 cluster (comprised of PVA OTUs 2, 3, and 4), which had a phylogeny contained in the *Crenarchaeota*. PVA OTU 1 was phylogenetically contained in the *Euryarchaeota*. Both of these lineages were phylogenetically affiliated with recently discovered cosmopolitan marine archaeoplankton. Chapter 5 represents a manuscript currently in preparation for submission under the authorship of C. L. Moyer, F. C. Dobbs, and D. M. Karl and was the basis for an abstract presented at the American Society for Microbiology 1995 general meeting (Moyer *et al.*, 1995b).

Finally, Chapter 6 provides a summary of the results from this dissertation project and suggests directions for future research.
REFERENCES


CHAPTER 2

A COMPUTER-SIMULATED RFLP ANALYSIS OF BACTERIAL SSU rRNA GENES: EFFICACY OF SELECTED TETRAMERIC RESTRICTION ENZYMES

ABSTRACT

An assessment of 10 tetrameric restriction enzymes (TREs) was conducted using a computer-simulated restriction fragment length polymorphism (RFLP) analysis for over 100 bacterial SSU rRNA gene sequences. The objective was to find the best combination of TREs for initial RFLP screening of cloned SSU rRNA genes. The resultant fragments were categorized based on the frequency of different band size classes. Three groups of distribution patterns for the TREs were determined via graphical exploratory data analysis. The RFLP data for each group of enzymes was then used to infer phylogenetic relationships with the neighbor-joining method, and the resulting bootstrap values and correct placement of node bifurcations were used as additional criteria for evaluation of the efficacy of the selected TREs. These computer-simulated RFLP data were compared to known phylogenetic relationships based on SSU rRNA sequence analysis. Of the 10 restriction enzymes examined, HhaI, RsaI and BstUI were determined to be the most efficacious at detecting and separating bacterial SSU rRNA genes based on their phylogeny (i.e., classifying operational taxonomic units or OTUs). These TREs are therefore recommended for use in future studies using SSU rRNA genes as descriptors of microbial diversity.
INTRODUCTION

A major objective of biological research in the past decade has been the detection and description of the microbial diversity present in the plethora of habitats found on Earth (e.g., Holben and Tiedje, 1988; Pace et al., 1986; Stahl et al., 1988). The development of the novel molecular biological techniques of polymerase chain reaction (PCR) and restriction fragment length polymorphism (RFLP) have made possible the rapid and precise identification and characterization of bacterial taxa from discrete habitats (Tiedje, 1995). A cornerstone method in this approach has been the use of PCR to amplify small-subunit (SSU) rRNA genes either from microbial communities or directly from novel culturable isolates. The stepwise procedure used to define in situ microbial diversity by estimating the in vivo cellular diversity present in a specific habitat using the in vitro genetic diversity of a sample is as follows: (i) efficiently isolate the native DNA, (ii) PCR-amplify SSU rRNA gene sequences with primers designed to react with group specific taxa, (iii) screen for genetic variability with an RFLP analysis, and (iv) use these detected variations to quantify genetic diversity and establish the phylogenetic relationships of novel isolates or clones. This study explores the efficacy of the screening for genetic variability process using a computer simulation to generate the restricted nucleotide fragments.

Recent studies of microbial diversity have screened mixed SSU rDNA clone libraries generated from bacterial communities using tetrameric (i.e., having a four-base pair recognition site) restriction enzymes (TREs) to identify putative operational taxonomic units (OTUs [DeLong et al., 1993; Haddad et al., 1995; Moyer et al., 1994]). This approach has
been successful, but the rationale for selecting a particular restriction enzyme or set of restriction enzymes has not yet been systematically evaluated. Normally, the choice of restriction enzymes has been based solely on practical matters such as cost and the probable number of fragments to be generated as determined by the size of the restriction site (Dowling et al., 1990; Moyer et al., 1994). So that the detection of sequence variations associated with bacterial SSU rRNA genes (SSU) could more readily and accurately occur, the efficacy of 10 TREs across a diverse spectrum of bacterial taxa using computer-simulated RFLPs was assessed. These taxa spanned the entire Bacteria domain (sensu Woese [Olsen et al., 1994; Winker and Woese, 1991; Woese, 1994; Woese et al., 1990]) and were used to test the hypothesis that the decision for TRE selection should be based on the enzyme's efficacy in distinguishing among known bacterial taxa. This hypothesis is based on a given TRE's site-specificity (i.e., sites of cleavage), which determines the frequency and distribution of the RFLPs produced from SSU rDNAs. It is this extrapolated site-specificity that ultimately dictated the selection of the most efficacious combination of TREs, to be recommended for use in the future screening of SSU rDNAs from unknown bacterial taxa.

Restriction fragments from the computer-simulated digestion of each TRE were categorized into a series of size classes (units = number of base pairs) and each bacterial taxon was determined to have either a presence or absence of a restriction fragment, or fragments, in each of these size class categories. These data were then analyzed using computer algorithms to estimate the phylogeny of the bacterial taxa that were the source for each of the representative SSU rDNAs and compared to similar phylogenetic
reconstructions based on the sequence data from each taxon’s SSU rDNA. These intercomparisons of the RFLP-based versus the DNA sequence-based phylogenetic reconstructions facilitated the assessment of the efficacy of each of the 10 TREs in differentiating among many (>100) bacterial taxa of known phylogeny (with respect to SSU rRNA sequences). Overall, this procedure allows for the comparison of a relatively rapid screening method using TREs with the more logistically-intensive, but precise standard method of DNA sequencing.

MATERIALS AND METHODS

Selection of bacterial sequence data included in model

A total of 106 bacterial SSU rDNA sequences were chosen from the Ribosomal Database Project (RDP [Maidak et al., 1994]), with the following additions: Alteromonas ANT-300 (C. L. Moyer, unpublished data), the representative hydrothermal vent clone isolates listed as PVB OTU’s (n=6 [Moyer et al., 1995]), Desulfitobacterium dehalogenans (Utkin et al., 1994), Azoarcus denitrificans (Zhou et al., 1995), and Desulfurella acetivorans (Rainey et al., 1993). This analysis focused on taxa spanning the domain Bacteria, including groups of taxa contained within each of the Proteobacteria subdivisions, and some taxa closely related at the generic level. Of the 106 bacterial sequences examined, 13 were contained in the α-Proteobacteria, 18 in the β-Proteobacteria, 23 in the γ-Proteobacteria, 15 in the δ- and ε-Proteobacteria, 16 in the Gram-positive phylum, and

23
21 distributed across several deeply-rooted phyla of Bacteria (See Appendix B for complete species names and phylogenetic affiliations of all taxa used in this study). Sequences were aligned based on primary and secondary structural considerations and were constructed using the GDE multiple sequence editor distributed through the RDP (Maidak et al., 1994). Computer-simulated RFLPs using the various TRE sites were conducted on the region of the SSU rDNA contained between the PCR priming sites described in Moyer et al. (1994), which yielded a fragment corresponding to positions 49 through 1,510 (Escherichia coli numbering system) for all the bacterial taxa examined. This region was chosen to give each bacterial taxon unambiguously alignable 5' and 3' SSU rDNA endpoints.

Computer generation of RFLPs

Sequence data from each bacterial taxon were restricted for each TRE using the program Mapsort, which is contained in the GCG version 7.3 sequence analysis software package. The TREs (with respective sequence recognition sites) used in the Mapsort program were as follows: AluI (AG'CT), BstUI (CG'CG), DdeI (C'TNAG), HaeIII (GG'CC), HhaI (GCG'C), Hinfi (G'ANTC), MboI ('GATC), MspI (C'CGG), Rsal (GT'AC), and TaqI (T'CGA). The output from Mapsort was tabulated and placed into a spreadsheet program (Quattro Pro for Windows, version 6.0). The presence or absence of SSU rDNA restriction fragments contained within size categories ranged as follows: 1,400 to 250 base pairs (bp) in 50 bp increments (± 25 bp precision), the 200 bp size category ranged from 225 to 187.5 bp, from 175 to 75 bp in 25 bp increments (± 12.5 bp precision),
and the 50 bp size category ranged from 62.5 to 45 bp. Due to the limitations of gel electrophoresis, 45 bp was judged as the lowest detectable limit. The frequency and distribution of restriction sites across all bacterial taxa were calculated directly from these RFLP data. For each TRE, the distribution of the size categories generated was represented graphically using an exploratory box-plot analysis (Tukey, 1977). The size range of ≥200 to ≤1000 bp was graphed as this range encompassed the majority of descriptive information. Such exploratory data analysis provided the basis for subsequent grouping of TREs according to the restriction fragment size-frequency distributions generated for each TRE.

Phylogenetic analyses of RFLP and DNA sequence data

All programs used in this study were taken from the PHYLIP version 3.5 phylogenetic analysis software package. Individual programs were SEQBOOT for data set bootstrapping, DNADIST for the calculation of evolutionary distances from DNA sequence data, NEIGHBOR for the neighbor-joining method of phylogenetic reconstruction, and CONSENSE for the calculation of a consensus tree. The program RESTDIST was used to compute distance matrix data from all restriction fragment size class data. PHYLIP and RESTDIST were provided by Dr. Joe Felsenstein; RESTDIST is planned to be released in a future version of the PHYLIP package. For the purpose of comparative phylogenetic analysis, DNA sequence data were limited to the comparison of highly to moderately conserved nucleotide positions that were unambiguously alignable in all sequences, corresponding to residues 101 to 183, 220 to 451, 482 to 839, 847 to 998, 1037 to 1130,
and 1143 to 1440 (E. coli numbering system). Distance matrix calculations of corrected evolutionary distances were generated by the programs DNADIST and RESTDIST, using the Kimura (1980) two-parameter model of sequence evolution. The algorithm of neighbor-joining (Saitou and Nei, 1987), which uses a step-wise approach of phylogenetic reconstruction rather than a heuristic search method, was used to compute the phylogeny for 100 independent bootstrapped data sets and a consensus tree was generated from each of these data sets for each final tree examined. Bootstrap values (Felsenstein, 1985) were categorized as \(\geq 50\) (phylogenetically correct for all nodes from all trees observed) with additional \(\geq 20\) values (correct in the majority of nodes observed, with only the correct nodes being represented). The sequence-based and RFLP-based phylogenetic trees for the full data set were used as “benchmarks” to compare sub-sets of data from discrete groups of TREs. These categorized bootstrap values were used as a criterion to assess the ability of each group of TREs to successfully differentiate among taxa contained in the domain Bacteria based on their representative SSU rDNAs.

RESULTS

The computer-simulated RFLPs for the SSU rDNAs from over 100 taxa from the Bacteria domain resulted in a restriction fragment size-frequency distribution for each of the 10 TREs tested (Fig. 2.1, 2.2, and 2.3). These data were analyzed by comparing each restriction fragment band size class generated for each TRE versus that size class’ overall band frequency across all taxa. This initial observation revealed three types of distribution
Figure 2.1. Restriction fragment size-frequency distributions cumulative across all taxa for the tetrameric restriction enzymes AluI, DdeI, MspI (Group 1). Horizontal dotted line shows threshold at 200 bp for restriction fragment band size classes. For restriction fragment band size classes from 250 to 1000 bp, shading of histogram bars changes at >20 and >40% band frequency.
Figure 2.2. Restriction fragment size-frequency distributions cumulative across all taxa for the tetrameric restriction enzymes *MboI, HinfI, TaqI, HaeIII* (Group 2). Horizontal dotted line shows threshold at 200 bp for restriction fragment band size classes. For restriction fragment band size classes from 250 to 1000 bp, shading of histogram bars changes at >20 and >40% band frequency.
The image shows four histograms labeled with different restriction enzymes:

- **Mbol**
- **TaqI**
- **Hinfl**
- **HaeIII**

Each histogram represents the frequency of band sizes in bp (base pairs) for each enzyme. The x-axis indicates the range of band sizes from 0 to 1000 bp, and the y-axis shows the band frequency percentage. The histograms are color-coded, with different shades indicating various frequency ranges.
Figure 2.3. Restriction fragment size-frequency distributions cumulative across all taxa for the tetrameric restriction enzymes HhaI, RsaI, BstUI (Group 3). Horizontal dotted line shows threshold at 200 bp for restriction fragment band size classes. For restriction fragment band size classes from 250 to 1000 bp, shading of histogram bars changes at >20 and >40% band frequency.
patterns for band size classes >200 to <1000 bp. This observation was further validated through an exploratory box-plot analysis (Tukey, 1977), which also demonstrated by inspection at least three types of distribution patterns (data not shown). The first distribution pattern (Group 1) was generated by the TREs AluI, DdeI, and MspI (Fig. 2.1). The Group 1 TRE distributions showed band frequencies in the 20 to 40% range for a majority of the size classes from 250 to 550 bp. There was only a single occurrence where a size class frequency reached above 40% in this group, which was the 250 bp size class for DdeI. The second type of distribution pattern (Group 2) resulted from the TREs MboI, Hinfl, TaqI, and HaeIII. The Group 2 distributions showed an increase in the occurrence of the larger band size classes ranging from 400 to 1,000 bp, but with none of these size classes present above a 20% band frequency (Fig. 2.2). In addition, only a single characteristic size class from between 250 to 350 bp obtained a band frequency above 25% (but in all cases reached above >40%) for each of the Group 2 TREs. The third type of distribution pattern (Group 3) was produced by the TREs HhaI, RsaI, and BstUI. This pattern showed an increased level of band frequency for the size classes ranging from 250 to 550 bp (Fig. 2.3). These band frequency levels were elevated to between 40 to 60% for at least two size classes and >20% for five size classes within the 250 to 550 bp range for each of the Group 3 TREs. The Group 3 distribution pattern displayed more >200 bp band size class occurrences (i.e., highest levels of band frequency) of all TREs examined.

The average number of cuts (restriction sites) per taxon for each TRE was estimated across all taxa (Table 2.1) along with the respective standard deviations. The result showed a general range of 3.71 to 5.65 average sites per taxon (across each SSU rDNA) for the 10
Table 2.1. Average number of restriction sites per taxon for each tetrameric restriction enzyme (TRE).

<table>
<thead>
<tr>
<th>TRE</th>
<th>TRE group$^a$</th>
<th>Mean number of restriction sites per taxon</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AluI</td>
<td>Group 1</td>
<td>4.54</td>
<td>1.19</td>
</tr>
<tr>
<td>DdeI</td>
<td>Group 1</td>
<td>4.24</td>
<td>0.98</td>
</tr>
<tr>
<td>MspI</td>
<td>Group 1</td>
<td>5.21</td>
<td>1.12</td>
</tr>
<tr>
<td>MboI</td>
<td>Group 2</td>
<td>3.97</td>
<td>1.34</td>
</tr>
<tr>
<td>HinII</td>
<td>Group 2</td>
<td>3.91</td>
<td>0.91</td>
</tr>
<tr>
<td>TaqI</td>
<td>Group 2</td>
<td>3.71</td>
<td>0.94</td>
</tr>
<tr>
<td>HaeIII</td>
<td>Group 2</td>
<td>4.80</td>
<td>1.21</td>
</tr>
<tr>
<td>HhaI</td>
<td>Group 3</td>
<td>4.17</td>
<td>1.09</td>
</tr>
<tr>
<td>RsaI</td>
<td>Group 3</td>
<td>4.48</td>
<td>0.93</td>
</tr>
<tr>
<td>BstUI</td>
<td>Group 3</td>
<td>5.65</td>
<td>1.12</td>
</tr>
<tr>
<td>All TREs</td>
<td>----</td>
<td>4.47</td>
<td>0.61</td>
</tr>
</tbody>
</table>

$^a$ As determined by the restriction fragment-size frequency distributions across all taxa.
TREs examined, with an overall average across all TREs of 4.47. The group mean number of cuts were 4.66, 4.10, and 4.77 sites per taxon for Groups 1, 2, and 3, respectively.

A phylogenetic tree was generated using SSU rDNA sequence data (Fig. 2.4) from all the taxa examined (see Appendix B for complete list of phylogenetically ordered taxa) and compared to multiple phylogenetic trees generated from RFLP data (data not shown). In all cases, the same neighbor-joining distance method was employed. As anticipated, the sequence-based phylogenetic tree generated in this study agreed with the accepted phylogenetic relationships for all representative taxa examined (based on SSU rDNA phylogeny). The sequence-based phylogenetic tree demonstrated that SSU rDNA sequence data were sufficient to yield bootstrap values at the majority of nodes at ≥50 and at all nodes at ≥20 (Fig. 2.4). This result was then compared with the RFLP-based phylogenetic tree that was constructed using the entire data set from all 10 TREs (Table 2.2). This “best-case scenario” RFLP-based phylogenetic tree demonstrated the generation of 25 bootstrap values ≥50 and an additional 23 bootstrap values ≥20. This RFLP-based phylogenetic tree was used as a metric by which to ascertain the success of the individual groups of restriction enzymes in predicting correct phylogenetic outcomes using RFLP data sub-sets from TRE Groups 1, 2, and 3 (Table 2.2).

The data from any single TRE or combination of two TREs were insufficient to generate reproducible bootstrap values or accurate phylogenetic relationships (data not shown). When the data from three TREs were combined from the RFLP analysis of SSU rDNAs, the results were sufficient to yield phylogenetic trees with reproducible bootstrap values at phylogenetically correct node bifurcations. This observation, together
Table 2.2. Number of phylogenetically estimated bootstrap values for each tetrameric restriction enzyme (TRE) group.

<table>
<thead>
<tr>
<th>Bootstrap Values</th>
<th>All TRE data</th>
<th>TRE Group 1</th>
<th>TRE Group 2</th>
<th>TRE Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 50</td>
<td>25</td>
<td>4</td>
<td>3</td>
<td>7&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>≥ 20</td>
<td>23</td>
<td>20</td>
<td>16</td>
<td>23</td>
</tr>
</tbody>
</table>

<sup>a</sup> Bold enhancement indicates highest number of occurrences among the three TRE groups.
Figure 2.4. Phylogenetic tree based on the SSU rDNA sequence data for all bacterial taxa examined (see Appendix B for phylogenetically ordered taxa list) using the neighbor-joining distance method. Bootstrap values ≥50 are shown in bold. Bootstrap values ≥20 are also shown. The scale bar represents 10 fixed mutations per 100 nucleotide positions.
with the initial RFLP distribution patterns was used as the basis for the further examination of the RFLP-based data through phylogenetic analysis.

The Group 1 tree (from the combined restriction data of AluI, DdeI, and MspI) yielded a total of 4 bootstrap values ≥50 and an additional 12 bootstrap values ≥20 (Table 2.2). The Group 2 tree was ultimately constructed without using the restriction data from HaeIII, as this TRE was the least phylogenetically informative of the group, having the highest average number of cuts per taxon (Table 2.1) and the greatest abundance of band size class frequencies <200 bp (Fig. 2.2). The premise that HaeIII was the least informative Group 2 TRE was confirmed by the addition of HaeIII restriction data and the sequential deletion of each of the other Group 2 TREs restriction data from the analysis, a process which in every case yielded lower numbers of descriptive bootstrap values (data not shown). Eliminating HaeIII was also necessary so that the final analysis would compare normalized RFLP-data sets, each combined from three TREs. The Group 2 phylogenetic tree (from the combined restriction data of MboI, HinfI, and TaqI) yielded a total of 3 bootstrap values ≥50 and an additional 16 bootstrap values ≥20 (Table 2.2). The Group 3 tree (from the combined restriction data of HhaI, RsaI, and BstUI) yielded a total of 7 bootstrap values ≥50 and an additional 23 bootstrap values ≥20 (Table 2.2). The TREs contained in Group 3 demonstrated the largest number of descriptive bootstrap values of the three groups of TREs tested.

The percent "successful" phylogenetic affiliations from the RFLP-based trees are presented in Table 2.3, as determined by phylogenetically correct node bifurcations with respect to SSU rRNAs (Fig. 2.4). The entire 10 TRE dataset was the most efficacious at
Table 2.3. Percent (%) “successful” phylogenetic affiliations across tetrameric restriction enzyme (TRE) groups as determined by phylogenetic analysis of RFLP data.

<table>
<thead>
<tr>
<th>Phylogenetic Group</th>
<th>All TRE data</th>
<th>TRE Group 1</th>
<th>TRE Group 2</th>
<th>TRE Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deeply-rooted phyla</td>
<td>76</td>
<td>52</td>
<td>52</td>
<td>67&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Gram-positive phylum</td>
<td>81</td>
<td>50</td>
<td>81</td>
<td>75</td>
</tr>
<tr>
<td>α-Proteobacteria</td>
<td>100</td>
<td>&lt;b&gt;100&lt;/b&gt;</td>
<td>77</td>
<td>92</td>
</tr>
<tr>
<td>β-Proteobacteria</td>
<td>94</td>
<td>78</td>
<td>94</td>
<td>83</td>
</tr>
<tr>
<td>γ-Proteobacteria</td>
<td>96</td>
<td>74</td>
<td>&lt;b&gt;83&lt;/b&gt;</td>
<td>&lt;b&gt;83&lt;/b&gt;</td>
</tr>
<tr>
<td>δ- and ε- Proteobacteria</td>
<td>100</td>
<td>53</td>
<td>47</td>
<td>&lt;b&gt;87&lt;/b&gt;</td>
</tr>
<tr>
<td>All taxa</td>
<td>91</td>
<td>67</td>
<td>73</td>
<td>80</td>
</tr>
</tbody>
</table>

<sup>a</sup> Bold enhancement indicates highest percentage value(s) among the three TRE groups.
describing phylogenetic relationships with respect to node bifurcations in addition to bootstrap values, as no TRE group yielded a greater percentage successful affiliations (i.e., 91% across all taxa). Of the three TRE groups, Group 3 was the most efficacious overall at predicting correct phylogenetic node bifurcations. Based on the phylogenetic groups examined (see Appendix B for full description), TRE Group 1 showed a greater percentage of successful affiliations for the $\alpha$-Proteobacteria at 81%. TRE Group 2 showed a greater percentage of successful affiliations for both the Gram-positive phylum and the $\beta$-Proteobacteria at 81 and 94%, respectively. TRE Group 3 was most successful in describing the phylogenetic affiliations for the deeply-rooted bacterial phyla and the $\delta$- and $\epsilon$-Proteobacteria (at 67 and 87%, respectively), in addition to being the most successful overall for all the taxa examined (80 versus 67 and 73% for TRE Groups 1 and 2, respectively).

DISCUSSION

RFLP analyses have been used extensively and successfully in the fields of systematic and evolutionary biology. The widespread use of this particular molecular biological technique provided the impetus behind the modeling effort reported herein. Restriction site variations have several appropriate characteristics, which allow for their use in both the differentiation of SSU rDNAs and phylogenetic reconstructions. The first and foremost is that RFLP data, though providing less direct information on the evolution of DNA sequences, are much cheaper and logistically easier to obtain. This is still the case today for
many laboratories, even with the current advances in DNA sequencing technology. Three characteristics that RFLP and sequencing methods have in common are: (i) character states can be scored unambigously, (ii) a large number of characters can be scored for each taxon, and (iii) information on both the extent and the nature of divergence between two DNA sequences are provided (Holsinger and Jansen, 1993). A critical characteristic that differs between RFLP and DNA sequence analyses relates to the asymmetry with respect to the evolution of restriction sites versus nucleotide positions. The 4 base pair recognition site for a TRE will be inactivated by any one of 12 different nucleotide substitutions, whereas if the sequence differs from the recognition site at only a single site, then only 1 substitution can occur to produce that restriction site. As a consequence, convergent losses of a restriction site are more likely than convergent gains, and the ratio of convergent losses to convergent gains increases as taxa become more divergent (Templeton, 1983). These characteristics must be considered when deciding which types of algorithms are best used to ascertain phylogenetic relationships with RFLP data, as reviewed by Holsinger and Jansen (1993).

This study has focused on specific criteria to evaluate TREs used for accurately detecting and differentiating among bacterial taxa (OTUs) based on their representative SSU rDNAs. The specific criteria were: (i) the analysis of the restriction fragment band size class frequency distributions, and (ii) the phylogenetic reliability of three groups of TREs based on their bootstrap values and phylogenetic affiliations. Other criteria may also be useful in deciding among TREs (e.g., the differentiation of closely related SSU rDNA clones). This latter criterion would best be satisfied using the TREs BstUI, MspI, and
HaeIII, based solely on their high average number of restriction sites per taxon (Table 2.1), which is also represented by a high frequency of <200 bp band size classes Fig. 2.1, 2.2, and 2.3).

The TREs contained in Group 3 (*HhaI*, *RsaI*, and *BstUI*) were superior for the differentiation of bacterial taxa, based on the criteria of RFLP distribution patterns and phylogenetic reliability for SSU rDNA sequences. This is most likely because Group 3 TREs yield the highest frequency (up to 60% across all taxa) of restriction fragment band size classes in the 250 to 550 bp range (Fig. 2.3). Of the Group 3 TREs, *HhaI* and *RsaI* are superior choices, because *BstUI* has a significantly greater average number of cuts per taxa (5.65 versus 4.17 and 4.48 for *HhaI* and *RsaI*, respectively; Table 2.1) and because *BstUI* has much higher frequency values for the band size classes of <200 bp (Fig. 2.3). It has been demonstrated that the simultaneous use of two TREs is sufficient to distinguish between SSU rDNAs (e.g., PVB OTUs 1 and 11) varying by a single *RsaI* site and with an overall total of seven variable nucleotide positions across the SSU rRNA gene (Moyer et al., 1994; Moyer et al., 1995). It is assumed that this level of variation is the approximate lower limit of resolution for this type of RFLP distribution analysis when used to detect OTUs. However, the primary conclusion of this study is that the best results for screening clone libraries of SSU rDNAs from unknown bacterial taxa would be through the individual use of the Group 3 TREs *HhaI*, *RsaI* and *BstUI* since these have been shown to be superior (relative to the 10 TREs tested) for accurately differentiating among many (>100) diverse bacterial taxa.

In a broader sense, this simulation has demonstrated that the choice of restriction
enzymes for screening and detecting bacterial taxa from community SSU rDNAs can and should be based on site-specificity. Furthermore, the frequency and distribution of the restriction enzyme site-specificities as modeled in this computer-simulated RFLP analysis have allowed for this selection. The goal of this exercise was to provide an efficient method, based on specific phylogenetic criteria, to explore microbial diversity from different habitats. Gaining a better understanding of microbial diversity is important to a sustainable biosphere, because individual species as well as communities of microorganisms play critical roles in the recycling of nutrients and minerals, production and consumption of gases that effect the global climate, pollutant degradation, and biological control of plant and animal pests. The exploration of such diversity is also a potential source of new genes and gene products for applications in biotechnology.
REFERENCES


CHAPTER 3

ESTIMATION OF DIVERSITY AND COMMUNITY STRUCTURE THROUGH RFLP DISTRIBUTION ANALYSIS OF BACTERIAL 16S rRNA GENES FROM A MICROBIAL MAT AT AN ACTIVE, HYDROTHERMAL VENT SYSTEM, LOIHI SEAMOUNT, HAWAI'I

ABSTRACT

PCR was used to amplify (eu)bacterial small-subunit (16S) rRNA genes from total-community genomic DNA. The source of total-community genomic DNA for this culture-independent analysis was the microbial mats from a deep-sea, hydrothermal vent system, Pele's Vents, located at Loihi Seamount, Hawaii. Oligonucleotides complementary to conserved regions in the 16S rRNA-encoding DNA (rDNA) of bacteria were used to direct the synthesis of PCR products, which were then subcloned by blunt-end ligation into phagemid vector pBluescript II. Restriction fragment length polymorphism patterns, created by using tandem tetrameric restriction endonucleases, revealed the presence of 12 groups of 16S rRNA genes representing discrete operational taxonomic units (OTUs). The rank order abundance of these putative OTUs was measured, and the two most abundant OTUs accounted for ~73% of the 16S rDNA clones. Among the remaining ~27% of the 16S rDNA clones, none of the 10 OTUs was represented by more than three individual clones. The cumulative OTU distribution for 48 bacterial 16S rDNA clones demonstrated
that the majority of taxa represented in the clone library were detected, a result which is assumed to be an estimate bacterial diversity in the native hydrothermal vent habitat. To confirm OTU specificity and 16S rDNA identity, 16S rDNA fingerprinting of individual clones belonging to particular OTUs was conducted using an oligonucleotide probe that binds to a universally conserved region of the 16S rDNA fragments.

INTRODUCTION

Defining the diversity and structure of natural microbial communities through the quantification of their constituent populations has been a long-standing challenge in microbial ecology. Selective enrichment cultivation as an approach for the description of naturally occurring microbial communities has severe limitations (Poindexter and Leadbetter, 1986; Ward et al., 1992), because the majority (typically >90 to 99%) of bacteria in nature cannot be cultured by traditional techniques (Brock, 1987; Jannasch and Jones, 1959). Consequently, it is unlikely that collections of bacterial isolates are representative of in situ diversity and community structure. Furthermore, because relatively nutrient-rich media are generally used for isolation, they may select for copiotrophic bacteria rather than dominant bacteria in the natural community.

Recently developed techniques in molecular biology provide for a culture-independent analysis of microorganisms and, therefore, an alternative approach to understanding the composition of natural communities. This approach examines variations in 16S rRNA or 16S rRNA-encoding DNA (rDNA) within naturally occurring prokaryotic
communities (Britschgi and Giovannoni, 1991; DeLong, 1992; DeLong et al., 1993; Fuhrman et al., 1992; Fuhrman et al., 1993; Giovannoni et al., 1990; Schmidt et al., 1991a; Ward et al., 1990a,b; Ward et al., 1992; Weller et al., 1991; Weller et al., 1992). The objective of the present study was to estimate diversity and community structure by performing an analysis of restriction fragment length polymorphism (RFLP) for all clones derived from a library of (eu)bacterial 16S rRNA genes following a tandem tetrameric site-specific restriction endonuclease treatment. The 16S rRNA gene contains information which makes it an excellent biomarker of microorganisms. For example, each 16S rRNA gene contains both highly conserved regions found among all living organisms and diagnostic variable regions unique to particular organisms or closely related groups. Analysis of the latter regions leads to a specific RFLP pattern, which can be used to define an operational taxonomic unit (OTU). In this study, tandem tetrameric restriction endonucleases were used to produce RFLP patterns, which were then analyzed and cataloged. Furthermore, 16S rDNA fingerprinting were used to validate the RFLP analysis results and the 16S rDNA identities.

"Community structure" is defined in terms of the number of OTUs present in a community and the abundance of individual clones within each OTU. These values were estimated by a tandem tetrameric RFLP analysis of each 16S rDNA clone. "Diversity" is a metric for the number of populations in a community and the genetic relatedness among these populations. This study focuses only on the populations in the bacterial community; genetic relatedness, as assessed by phylogenetic analysis, will be discussed in a subsequent dissertation chapter. For both community structure and diversity, the number of populations
was estimated by the number of OTUs present in the community. In addition, an analog of rarefaction was used to ascertain that the majority of the total diversity in the native habitat was successfully detected with the 48 clones examined.

The microbial community and habitat examined were the microbial mats at Pele's Vents, which are located on the southwest portion of the summit (980 m below sea level) of Loihi Seamount. The area of active hydrothermal fluid discharge is restricted to the flank of a relatively small volcanic cone approximately 10 to 15 m below the summit. The active field (area, <0.25 km²) is characterized by numerous individual vents discharging waters heated to temperatures of ≤37°C (compared with the ambient seawater temperature of 4°C). The individual vent orifices are distinguished by a white precipitate, which has been determined to contain high concentrations of elemental sulfur (Karl et al., 1988; Karl et al., 1989). The vent fluid is exceptionally clear and nearly devoid of suspended particulate matter.

Loihi Seamount is an active, submarine, hotspot volcano and potentially the next Hawaiian island. Geochemical and biological evidence suggest that hydrothermal vent systems at mid-plate sites differ fundamentally from vent systems at plate boundaries (e.g., mid-ocean ridges) (Karl et al., 1988; Karl et al., 1989; Sedwick et al., 1992). Most notably, the vent waters at Pele's Vents contain extremely high concentrations of total dissolved CO₂ (ca. 300 mM), more than 100 times greater than the concentration at the Galapagos Rift Vents (Edmond et al., 1987). Consequently, in situ pH of vent waters can be as low as 4.2 (Sedwick et al., 1992). The levels of dissolved iron (ca. 1 mM) are 2 × 10⁶ times greater than the levels in ambient seawater and approximately 40 times greater than the levels at the
Galapagos Rift Vents (Edmond et al., 1987). Another compound of potential microbial importance is CH₄; the dissolved CH₄ concentration (ca. 7.2 μM) in the effluent of Pele's Vents is enriched 10⁵-fold compared with ambient seawater (Karl et al., 1989). Loihi Seamount has none of the luxuriant macrobenthos (e.g., tubeworms and giant clams that harbor bacterial endosymbionts) present at other deep-sea hydrothermal systems. This characteristic may be related to the absence or extremely low levels of sulfide in the vent waters (Sedwick et al., 1992), to the high levels of dissolved CH₄ and iron (Karl et al., 1988; Karl et al., 1989), or perhaps to the low pH of the effluent waters (Sedwick et al., 1992).

There have been previous descriptive and metabolic studies of the microbiological components at Loihi Seamount. Karl et al. (1988, 1989) described the predominance of iron-depositing sheathed bacteria at the site, demonstrated methanotrophy, utilization of acetate and glutamate, and incorporation of [³H]adenine and [³H]thymidine into nucleic acids, and detected thermophilic bacterial populations. This study supplements this initial information and estimates the diversity and community structure of bacteria, as previously defined, at Pele's Vents.

**MATERIALS AND METHODS**

*Collection of bacterial mat samples*

Material was collected from Pele's Vents (depth, 980 m) on dives 193 to 196 (12 October 1991 to 18 October 1991) and 208 to 215 (18 September 1992 to 1 October 1992)
of DSRV Pisces V (Fig. 3.1). Two independent methods were used to sample the microbial mats. Mat-covered rocks were placed into polyvinylchloride "coffins" to prevent winnowing of the mat material en route to the surface. Hydrothermal fluids containing bacterial mat material were collected by using the vacuum produced by opening 2 liter Niskin baggie samplers adapted with a 2 m Tygon tube for directional sampling. Aboard ship, samples were collected either by aseptically scraping mat-covered rocks or by allowing mat particulates to settle inside suspended Niskin baggies. Samples were quick frozen and maintained on dry ice until they were returned to the laboratory, and then they were kept at -85°C until further processing.

*Genomic DNA extraction and purification*

Both French press lysis and enzymatic cellular lysis were used to maximize the recovery of genomic DNA. Approximately 10 g (wet weight) of mat material was thawed on ice, ground with a cold (4°C) sterile mortar and pestle, and resuspended in cold (4°C) sucrose lysis buffer (0.75 M sucrose, 700 mM NaCl, 40 mM Na₂EDTA, 50 mM Tris-HCl; pH 8.3) to a total volume of 20 ml; the high salt concentration helped prevent DNA shearing. The mat slurry was passed through a cold (4°C) French pressure cell twice at 20,000 lb/in², 1 mg of lysozyme per ml was added, and then the preparation was incubated at 37°C for 30 min. After the addition of 0.5% (wt/vol) sodium dodecyl sulfate (SDS), 100 μg of proteinase K (United States Biochemical Co., Cleveland, OH.) per ml, 250 μg of achromopeptidase (Sigma, St. Louis, MO.) per ml, and 50 μg of RNase A (United States
Figure 3.1. Map showing location of Loihi Seamount in relation to the Big Island of Hawaii. (Inset) Larger-scale view of the major islands which comprise the state of Hawaii. Redrawn from Fornari et al. (1988). Depth contours (isobaths) are in meters.
Biochemical Co.) per ml, the mixture was incubated at 55°C for 30 min. The cells in the mat slurry were monitored by microscopic examination to determine that complete lysis had occurred. The polysaccharides and residual proteins were aggregated by adding 1% (wt/vol) hexadecyltrimethyl ammonium bromide (CTAB) and incubating the preparation at 65°C for 30 min. Protein and polysaccharide complexes were removed by extraction with an equal volume of phenol-chloroform-isoamyl alcohol (50:49:1) (phenol was first prepared fresh by water-saturation, 0.1% [wt/vol] p-hydroxyquinoline addition, and buffer saturation with STE buffer [100 mM NaCl, 10 mM Na₂EDTA, 50 mM Tris-HCl; pH 7.4]). The residual phenol was removed by extracting the aqueous phase with an equal volume of chloroform-isoamyl alcohol (24:1). The genomic DNA was allowed to precipitate at -20°C for 8 h after the addition of 0.1 volume 3 M sodium acetate (pH 4.6) and 2.5 volumes of 100% ethanol. It was then collected by centrifugation at 10,000 x g for 30 min using 30-ml Corex glass centrifuge tubes (Corning Inc., Horseheads, N.Y.). The genomic DNA pellet was resuspended to a concentration of 100 ng/µl, as measured A₂₆₀, and the yield was between 10 and 20 µg/10 g (wet weight) of bacterial mat sample. Approximately 5 to 10 µg of this crude DNA preparation was purified by treatment with Qiaex (Qiagen Inc., Chatsworth, CA.), a uniform 3 µm silica gel matrix which selectively binds DNA in the presence of high salt concentrations.

Oligonucleotide synthesis and purification

All oligonucleotides used in this study were synthesized with a model 380B
automated DNA synthesizer (Applied Biosystems, Foster City, CA.) and purified using the Surepure thin-layer chromatography system (United States Biochemical Co.) at the University of Hawaii Biotechnology-Molecular Biology Instrument and Training Facility.

Amplification of 16S rRNA genes

The 16S rDNA was selectively amplified from purified genomic DNA using PCR (Erlich et al., 1991; Saiki et al., 1988; Steffan and Atlas, 1991) with oligonucleotide primers designed to anneal to conserved positions in the 3' and 5' regions of bacterial 16S rRNA genes. The forward primer (5'-TNANACATGCAAGTCGAICG) corresponded to positions 49 to 68 of *Escherichia coli* 16S rRNA (Brosius et al., 1978), and the reverse primer (5'-GGYTACCTTGTTACGACTT) corresponded to the complement of positions 1,510 to 1,492 (Lane, 1991). The reaction conditions were as follows: 100 ng of template DNA, 10 µl of 10× PCR reaction buffer (500 mM KCl, 100 mM Tris-HCl [pH 9.0 at 25°C], 15 mM MgCl₂, 1% [wt/vol] Triton X-100), 2.5 U of Taq DNA polymerase (Promega, Madison, WI.), 1.5 µg of T4 gene 32 protein (United States Biochemical Co.), 1 µM forward primer, 1 µM reverse primer, 200 µM dATP, 200 µM dCTP, 200 µM dGTP, and 200 µM dTTP were combined in a total volume of 100 µl. The reaction mixtures lacking template DNA, T4 gene 32 protein, and Taq DNA polymerase were UV irradiated for 10 min to eliminate potentially contaminating template DNA (Sarkar and Sommer, 1990) and then heated at 94°C for 2 min. The complete reaction mixture was overlaid with mineral oil and incubated in a thermal cycler (Perkin-Elmer Cetus, Norwalk, CT.) as
follows: denaturation at 94°C for 1 min, primer annealing at 60°C for 1.5 min, and chain extension at 72°C for 3 min with an additional extension time of 7 min on the final cycle, for a total of 33 cycles. Each lot of T4 gene 32 protein and Taq DNA polymerase was tested for potential contamination of 16S rDNA templates as described by Schmidt et al. (1991b). PCR-amplified gene products were separated chromatographically in 1.5% agarose (0.75% NuSieve and 0.75% SeaKem; FMC Bioproducts, Rockland, ME.) electrophoresis gels stained with 0.5 µg of ethidium bromide per ml and visualized by UV excitation.

Construction of bacterial 16S rDNA clone library

Amplified 16S rDNA gene products from four to six individual PCRs were pooled, purified by treatment with Qiaex (Qiagen Inc.), and made blunt ended by treatment with 10 U of the large (Klenow) fragment of DNA polymerase I and 10 U of T4 polynucleotide kinase (Promega). The reaction mixture also included 10 µl of 10x Klenow buffer (0.5 M Tris-HCl [pH 7.5 at 25°C], 100 mM MgCl₂, 10 mM dithiothreitol, 0.5 mg of bovine serum albumin per ml), 1 mM ATP, 200 µM dATP, 200 µM dCTP, 200 µM dGTP, and 200 µM dTTP; the total volume was 100 µl, and the preparation was incubated at 37°C for 1.0 h. The blunt-ended PCR-amplified 16S rDNA gene products were again purified by treatment with Qiaex and ligated into the SmaI restriction site of pBluescript II KS⁺ (Stratagene, La Jolla, CA.). The ligation reaction mixture contained 270 ng of insert and 15 ng of vector (i.e., the molar ratio was 9:1), as well as 1 µl of 10x ligase buffer (300 mM Tris-HCl
[pH 7.8 at 25°C], 100 mM MgCl₂, 100 mM dithiothreitol, 10 mM ATP) and 2 U of T4 DNA ligase (Promega) in a final reaction volume of 10 µl. The mixture was incubated at 16°C for 12 to 14 h. The resulting ligation products from four independent ligation reactions were pooled, diluted 2:3 with TE buffer (10 mM Tris-Cl, 1 mM Na₂EDTA [pH 8.0 at 25°C]), and used to transform competent Epicurian Coli XL1-Blue MRF cells (Stratagene) according to the manufacturer's protocol. Clones were screened for α-complementation by using X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) as the substrate (Sambrook et al., 1989) on YT agar supplemented with ampicillin (100 µg/ml) and tetracycline (12.5 µg/ml).

16S rDNA RFLP analysis

Plasmid DNA was prepared from clones by the Magic Miniprep DNA purification system (Promega). Insert 16S rDNA gene fragments were cut from the plasmid vector with restriction enzymes BamHI plus PstI (United States Biochemical Co.), concentrated by ethanol precipitation with ammonium acetate (Sambrook et al., 1989), separated by gel electrophoresis in 1.5% agarose (0.75% NuSieve and 0.75% SeaKem; FMC Bioproduots) gel, stained with 0.5 µg of ethidium bromide per ml, and visualized by UV excitation. Insert 16S rDNA gene fragments were excised from the agarose gel, extracted by using Qiaex, restricted by using the tetrmeric endonuclease pairs of HaeIII plus MspI or AluI plus Rsal (New England Biolabs Inc., Beverly, MA.), and concentrated by ethanol precipitation with ammonium acetate (Sambrook et al., 1989). The resulting RFLP products were separated
by gel electrophoresis in 4.0% agarose (3.0% NuSieve and 1.0% SeaKem; FMC Bioproducts) gels stained with 0.5 μg of ethidium bromide per ml and visualized by UV excitation.

16S rDNA fingerprinting by Southern blotting and oligonucleotide probe hybridization

RFLP gels were denatured in a 1.5 M NaCl–0.5 N NaOH solution and Southern blotted (Southern, 1975) onto Nytran membranes (Schleicher & Schuell, Keene, N.H.) with a vacuum blotter (Model 785; Bio-Rad Laboratories, Richmond, CA.). The transfer buffer was 2.16 M NaCl, 12 mM Na₂EDTA, 120 mM NaPO₄ buffer (pH 7.6) (Grimont and Grimont, 1991). The following DNA oligonucleotides were used as hybridization probes with RFLP blots: oligonucleotide 1406F (5'-GYACACACCGCCCGT), corresponding to E. coli 16S rRNA positions 1,392 to 1,406; oligonucleotide 926F (5'-CAGCMGCCGCGGTAATIC) corresponding to positions 907 to 926; and oligonucleotide 536F (5'-AAACTYAAAKGAATTGACGG) corresponding to positions 519 to 536 (Lane et al., 1985). These oligonucleotides correspond to highly conserved or "universal" regions found in all known small-subunit ribosomal genes (Lane et al., 1985). Each was 5' end labeled with [γ-³²P]ATP (Amersham Corp., Arlington Heights, IL.) using T4 polynucleotide kinase (Promega) to a specific activity of approximately 2 to 4 × 10⁸ dpm/μg (Stahl and Amann, 1991). The labeled probes were purified on C₁₈ reverse-phase Sep-Pak columns (Millipore Corp., Milford, MA) (Lane et al., 1985). Southern-blotted RFLP membranes were dried at 80°C for 30 min under a vacuum and UV cross-linked for 2 min (0.1 J/cm²).
The membranes were prehybridized for 15 min at 25°C in 10 ml of hybridization buffer (1× Denhardt's reagent, 0.1% SDS, 1.08 M NaCl, 6 mM Na₂EDTA, 60 mM NaPO₄ [pH 7.6]) in a hybridizer oven (Model HB-1D; Techne Inc., Princeton, N.J.). The solution was replaced with 10 ml of fresh hybridization buffer containing 20 μCi of [γ-³²P]-labeled oligonucleotide probe, and the preparation was incubated at 25°C for 4 to 12 h. Following hybridization, the membranes were washed three times (15 min each) in 25 ml of washing buffer (0.1% SDS, 1.08 M NaCl, 6 mM Na₂EDTA, 60 mM NaPO₄ [pH 7.6]) at 28°C. The final 15-min wash at 37°C was done in washing buffer. Hybridized membranes were kept moist for the purpose of reprobing by placing them on blotting paper saturated with TE buffer and sealed inside Micro-Seal bags (Dazey Corp., Kansas City, KS.). The Micro-Seal bags were placed directly into autoradiography cassettes with intensifying screens for 6 to 8 h of exposure to X-ray film. Oligonucleotide probes were stripped from membranes by washing the membranes twice (15 min each) with washing buffer at 65°C prior to reprobing with a different oligonucleotide.

RESULTS

A total of 76 colonies were chosen after they tested positive for α-complementation of β-galactosidase, and 51 clones contained an insertion detectable by primary restriction with BamHI plus PstI. A total of 48 of these clones contained the entire 1.5 kb 16S rDNA insert, including three discrete, universally conserved oligonucleotide hybridization sites. These universally conserved sites were assessed by 16S rDNA fingerprinting analyses (data
not shown). The RFLP patterns of the 48 intact 16S rDNA clones indicated that only three specific patterns were obtained after primary restriction with BamHI plus PstI (Table 3.1). These RFLP patterns resulted from no more than one internal cut site within the 16S rDNA insert of any individual clone. Two of the three cloned inserts not used in the OTU analysis contained incomplete inserts that were approximately 600 bp long, and the other clone had an approximately 1.8 kb insert containing a 16S rDNA chimeric structure. The latter was detected after the cloned 16S rDNA insert was digested after secondary restriction with tetrameric endonuclease pairs, which was followed by the hybridization of two discrete sites to a single universally conserved oligonucleotide probe (data not shown).

Secondary restriction of complete 16S rDNA inserts was performed with tetrameric endonuclease pairs (either HaeIII plus MspI or AluI plus Rsal) to identify discrete OTUs. In evaluating the RFLP patterns that emerged, each discrete pattern was classified as an OTU, which was either unique for a single clone or similar for two or more clones. After secondary restriction of the entire 16S rDNA clone library with HaeIII plus MspI, 11 OTUs were detected. These included all of the clones contained in the two dominant OTUs (data not shown). Secondary restriction of all 16S rDNA clones with AluI plus Rsal was slightly more discriminating in that one additional OTU, containing a single clone, was detected. OTU 11 was separated from OTU 1 in this manner. Overall, when AluI plus Rsal were used after excision with BamHI plus PstI, a total of 12 OTUs were detected (Table 3.1; Fig. 3.2).

The validity of the 12 OTUs was confirmed from the results of 16S rDNA fingerprinting of representative 16S rDNA clones that underwent secondary restriction with AluI plus Rsal using oligonucleotide probe 536F (Table 3.1; Fig. 3.3). In addition,
### TABLE 3.1. Data matrix for OTU polymorphisms as detected by restriction digestion

<table>
<thead>
<tr>
<th>Restriction digestion</th>
<th>Estimated size of rDNA fragment (bp)</th>
<th>OTU 1 ((n = 12))</th>
<th>OTU 2 ((n = 23))</th>
<th>OTU 3 ((n = 3))</th>
<th>OTU 4 ((n = 1))</th>
<th>OTU 5 ((n = 1))</th>
<th>OTU 6 ((n = 1))</th>
<th>OTU 7 ((n = 1))</th>
<th>OTU 8 ((n = 1))</th>
<th>OTU 9 ((n = 1))</th>
<th>OTU 10 ((n = 1))</th>
<th>OTU 11 ((n = 1))</th>
<th>OTU 12 ((n = 1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary (b)</td>
<td>1500 bp</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>1250 bp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>800 bp</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>700 bp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>250 bp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Secondary (c)</td>
<td>550 bp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>350 bp</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>300 bp</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>250 bp</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>4(d)</td>
</tr>
<tr>
<td></td>
<td>175 bp</td>
<td>+(d)</td>
<td>+</td>
<td>+(d)</td>
<td>+</td>
<td>+</td>
<td>+(d)</td>
<td>+</td>
<td>+</td>
<td>+(d)</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150 bp</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>125 bp</td>
<td>-</td>
<td>+(d)</td>
<td>+(d)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+(d)</td>
<td>+(d)</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>100 bp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>75 bp</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>4(d)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>50 bp</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+(d)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

\(n\) is number of clones examined.

\(b\) Restriction with BamHI plus PstI.

\(c\) Restriction with Alul plus RsaI.

\(d\) After hybridization with universal oligonucleotide probe 536F.
Figure 3.2. OTUs determined from RFLP analysis of representative cloned inserts of bacterial 16S rDNA restricted with AluI plus RsaI. Lanes A and J contain HaeIII-restricted \( \phi X174 \) DNA size markers. Lanes B, K, and R, OTU 1 clones; lanes C and L, OTU 2 clones; lanes D and M, OTU 3 clones; lane E, OTU 4 clone; lane F, OTU 5 clone; lane G, OTU 6 clone; lane H, OTU 7 clone; lane I, OTU 8 clone; lane N, OTU 9 clone; lane O, OTU 10 clone; lane P, OTU 11 clone; lane Q, OTU 12 clone. Two or three lanes containing members of the same OTU demonstrate that there was reproducibility among different 16S rDNA clones in that OTU.
Figure 3.3. 16S rDNA fingerprinting of RFLPs from clones representing OTUs with oligonucleotide probe 536F. Lanes A, I, and P, OTU 1 clones; lanes B and J, OTU 2 clones; lanes C and K, OTU 3 clones; lane D, OTU 4 clone; lane E, OTU 5 clone; lane F, OTU 6 clone; lane G, OTU 7 clone; lane H, OTU 8 clone; lane L, OTU 9 clone; lane M, OTU 10 clone; lane N, OTU 11 clone; lane O, OTU 12 clone. Two or three lanes containing members of the same OTU demonstrate that there was reproducibility among different 16S rDNA clones in that OTU.
16S rDNA fingerprinting confirmed that cloned inserts were intact 16S rDNA gene fragments; each of the three oligonucleotide probes hybridized to the universally conserved regions in the 16S rRNA gene (data not shown for probes 926F and 1406F). Intact 16S rDNA inserts were determined both from the overall sizes of cloned inserts after primary restriction with BamHI plus PstI from the phagemid vector and from the occurrence of a single hybridization site among the RFLP bands for each of the three universally conserved oligonucleotide probes.

Having identified 12 putative OTUs, the distribution of the 48 16S rDNA clones among these OTUs was determined (Fig. 3.4). The OTUs were numbered solely in the order of detection, which was assumed to be stochastic, as determined by the first 16S rDNA clone found in each OTU. OTUs 1 and 2 together accounted for ~73% of all of the 16S rDNA clones. Of the two dominant OTUs, OTU 2 clones were nearly twice as prevalent as OTU 1 clones; OTUs 1 and 2 accounted for ~25 and ~48%, of the 16S rDNA clones examined, respectively. The remaining 13 bacterial 16S rDNA clones were distributed among 10 OTUs. Only 2 of the remaining 11 OTUs contained more than one clone. Thus, 8 of the 12 OTUs were represented by a single 16S rDNA clone.

To determine whether in situ bacterial diversity was well described by the 16S rDNA clones examined, the cumulative number of OTUs was plotted as a function of clone number (Fig. 3.5). The bacterial 16S rDNA clones were numbered solely on the basis of initial detection, which was assumed to be stochastic. This technique is analogous to generating a rarefaction curve to estimate species richness from a deterministic transform of species abundance data (Tipper, 1979). After the first 27 bacterial 16S rDNA clones were
Figure 3.4. Distribution in OTUs of 48 bacterial 16S rDNA clones from the bacterial mat community at an active, hydrothermal vent system, Loihi Seamount, Hawaii. Abundance, as determined by the number of 16S rDNA clones found in each OTU, was used to define the community structure. The OTUs are shown in the order of initial detection.
Figure 3.5. Estimation of diversity in the bacterial mat community at an active, hydrothermal vent system, Loihi Seamount, Hawaii. The sequential detection of cumulative OTUs following RFLP analysis of a 48-clone bacterial 16S rDNA clone library is represented. For example, after 27 (and 32) clones were examined, 11 (and 12) OTUs were detected. The 16S rDNA clone numbers reflect the order of initial detection, which was assumed to be stochastic relative to the distribution of clones generated in the library.
examined, 11 of the total 12 OTUs had been detected. Only one additional OTU was detected among the remaining 21 clones. The two dominant OTUs were detected prior to any other OTUs and only after the first three 16S rDNA clones were examined.

**DISCUSSION**

This chapter demonstrates the novel use of tandem tetrameric restriction endonucleases with cloned 16S rDNA gene fragments to generate RFLP data (equivalent to ribotyping) combined with rDNA fingerprinting for the purpose of describing bacterial community structure and diversity. The focus of the study was a deep-sea, hydrothermal vent microbial mat community, and it is believed that this analytical strategy should apply equally well in other habitats. In general terms, a microbial community is defined as an assemblage of co-occurring microorganisms interacting at given location or habitat; it is the highest biological unit comprised exclusively of individuals and populations. Each population within a community has a distinct functional role or niche. There are a finite number of niches within a community, and these are filled by the intrinsic populations of that community (Atlas and Bartha, 1993). The term “community structure” encompasses the number of populations within a community and the number of individuals within each population. A community’s “diversity” describes the number of individual populations, as well as the relative genetic relatedness among these populations. Obtaining a better understanding of bacterial community structure and diversity is crucial to aspects of microbial ecology where bacteria interact with one another and with their environment (e.g.,
global biogeochemical cycling of matter, risk assessment related to the release of genetically engineered microorganisms, predator-prey relationships, and trophic-level interactions).

The use of tandem tetrameric restriction enzymes to establish OTUs was a keystone of the present approach. While any one of the four tetrameric restriction enzymes by itself yielded too few recognition sites for OTU analysis (data not shown), either pair of enzymes (HaeIII plus MspI or AluI plus RsaI) would have been sufficient for the detection of the majority of the OTUs present in the hydrothermal bacterial community examined. The 1.5 kb cloned 16S rDNA inserts yielded restriction fragments in the size range from 50 bp (lower limit of detection) to 800 bp. In a perfectly random DNA sequence containing 50% G+C, a tetrameric recognition site would occur every 256 bases. Therefore, two tetrameric restriction enzymes theoretically would yield 11 or 12 recognition sites within any given 1.5 kb gene fragment. Because of the moderately higher G+C contents of 16S rRNA genes (most bacterial 16S rDNA have G+C contents between 55 to 65% [Woese, 1987]), the restriction enzymes HaeIII plus MspI were initially used. However, AluI plus RsaI yielded more restriction fragments within the required size range and were slightly more discriminating for the bacterial 16S rDNA clones examined in this study (Table 3.1; Fig. 3.2), which resulted in the detection of an additional OTU.

An integral part of the analysis was rDNA fingerprinting through Southern blotting and 16S rDNA universal oligonucleotide probe hybridizations (Table 3.1; Fig. 3.3). This procedure confirmed the occurrence of 12 OTUs and the identities of the 16S rDNA inserts for each clone. In addition, a chimeric 16S rDNA clone most likely produced by PCR-mediated amplification was detected. Without either rDNA fingerprinting or a complete
secondary structure analysis of the primary sequence data from the entire 16S rRNA gene, this PCR-mediated potential risk could produce erroneous results and thereby suggest the presence of organisms that in fact do not exist (Liesack et al., 1991). The chimeric 16S rDNA clone was detected herein, after secondary restriction by hybridization with a universal oligonucleotide probe. This probing revealed the presence of dual sites for a universally conserved position that normally occurs only once in a 16S rDNA gene. This result, found during an initial inspection of the bacterial clone library by rDNA fingerprinting, causing the elimination of this suspect clone from the final OTU analysis.

The bacterial mat community collected at Pele's Vents was dominated by two OTUs, which accounted for ~73% of the total 16S rDNA bacterial clone library (48 clones). An additional 10 OTUs were present at low levels. The bacterial community structure, consisting of 12 OTUs detected among the 48 bacterial 16S rDNA clones and their respective levels of abundance, approximates a log-normal distribution (Fig. 3.4). This distribution of individual 16S rDNA clones in OTUs (community structure) for Pele's Vents is especially interesting as this habitat may be considered an "extreme" environment because of its low pH (~4.0), strong thermal gradient, high dissolved CO₂ (~300 mM), and elevated trace metal concentrations (e.g., an Fe concentration of ≥1 mM). This finding is accentuated by the ephemeral nature of hydrothermal vent water emissions in general, which directly affect the temperature of the habitat.

In examining the cumulative OTU distribution of 48 bacterial 16S rDNA clones (Fig. 3.5), the significant decrease in the rate of OTU cumulation was interpreted as evidence that most of the diversity in the clone library was detected by the RFLP analysis.
Upon examination, 11 of 12 OTUs among the first 27 bacterial 16S rDNA clones were detected. The remaining 21 bacterial 16S rDNA clones yielded only 1 additional OTU. As there certainly must be rarely occurring bacteria in the mat community, an examination of additional 16S rDNA clones may have detected more OTUs. It is clear, however, that the level of analysis was sufficient to detect the community's predominant OTUs and infer their distribution within the microbial mat community at Pele's Vents.

A crucial factor to address in an analysis of microbial community structure and diversity when 16S rRNA biomarkers are used is whether the focus of study should be on the potential genetic diversity represented at the DNA level or on physiologically active genetic diversity represented at the RNA level. No habitat can optimally support the growth of all of the bacteria that it contains (Morita, 1982; Morita, 1985), and in marine environments it is likely that starvation-survival processes are a common metabolic strategy for a majority of bacteria (Morita, 1986; Morita, 1988; Morita and Moyer, 1989). When in a starvation-survival state, marine bacteria lose viability as well as cellular DNA and RNA disproportionately, depending on the prestarvation growth rate (Morita and Moyer, 1989; Moyer and Morita, 1989a,b; Moyer et al., 1990). The rRNA content decreases predictably during starvation-survival, with loss rates that depend on the physiological state at the onset of starvation (Flärdh et al., 1992; Kramer and Singleton, 1992). A positive correlation between cellular ribosomal (rRNA) content over a wide range of growth rates for bacteria is a long-standing axiom in microbial physiology (Dennis and Bremer, 1974; Ingraham et al., 1983; Rosset et al., 1966; Schaechter et al., 1958). Recently, this relationship was demonstrated elegantly using fluorescent 16S rRNA-targeted hybridization probes for single
cells of \textit{E. coli} strain B/r (DeLong \textit{et al.}, 1989) and for single cells of sulfate reducers found in biofilms (Poulsen \textit{et al.}, 1993). Direct analysis of 16S rRNA can potentially bias diversity estimations in favor of rapidly growing populations of cells and can underestimate the genetic diversity present in a given habitat. Therefore, the 16S rRNA of a natural microbial community better estimates the physiologically active microorganisms than the absolute genetic diversity and community structure. Consequently, estimating diversity at the DNA level, rather than at the RNA level, theoretically provides a more accurate measurement of taxonomic group variability by potentially detecting slow growing or dormant microorganisms present within the community.

Estimating community structure and diversity at the DNA level is an invaluable tool for microbial ecology, but this strategy also has potential problems and limitations. The oligonucleotide primers used to amplify the 16S rRNA genes from the bacterial mat community are complementary to regions conserved over the entire bacterial domain. Therefore, it is assumed that the clone library contained an array of 16S rDNA approximately as diverse as that of the bacterial mat community at Pele's Vents. Furthermore, the distribution of 16S rDNA clones within a library ultimately should approximate the relative distribution of cells in the habitat. However, the possibility of selection during the DNA extraction process exists, and care must be taken to achieve the highest possible level of efficiency in cell extraction and DNA recovery to avoid selection prior to PCR amplification of the cloned 16S rDNA. It is also possible that "cell-free" detrital DNA, which may have been adsorbed onto the mineral-rich microbial mat material found at Pele's Vents was extracted. The potential for bias at the level of the PCR and
ligation reactions also exists, which is why care was taken to use multiple PCR and ligation reactions to construct the bacterial clone library. Finally, as shown previously, when conducting community structure and diversity analyses with these techniques, workers must be especially alert to and test for the possibility of PCR-mediated chimeric gene amplification.

In summary, the present study demonstrates a novel approach for estimating microbial diversity and community structure from environmental samples using recently developed molecular biological techniques. Tetrameric restriction endonuclease pairs were used to detect OTUs by an RFLP analysis of a PCR-amplified 16S rDNA bacterial clone library. Coupling this technique with rDNA fingerprinting, estimates the number of OTUs and the abundance of each OTU. Through the application of this approach to the bacterial mats at Pele's Vents, a deep-sea hydrothermal vent system, the bacterial community was shown to be dominated by 2 OTUs and contained at least 12 OTUs, entities analogous to bacterial species. Phylogenetic analyses of the 16S rRNA genes from each of the bacterial OTUs as well as archaeal community structure and diversity analyses from the microbial mats at Pele's Vents, the subject of latter chapters.
REFERENCES


CHAPTER 4

PHYLOGENETIC DIVERSITY OF THE BACTERIAL COMMUNITY FROM A MICROBIAL MAT AT AN ACTIVE, HYDROTHERMAL VENT SYSTEM, LOIHI SEAMOUNT, HAWAII

ABSTRACT

The phylogenetic diversity of small-subunit (SSU) rRNA genes associated with the domain Bacteria was examined for a microbial mat at an active, deep-sea hydrothermal vent system (using previously defined operational taxonomic units [C.L. Moyer, F.C. Dobbs, and D.M. Karl, Appl. Environ. Microbiol. 60:871-879, 1994]; those for Pele's Vents Bacteria are hereafter abbreviated PVB OTUs). A cluster of phylogenetically related PVB OTUs (2, 3, 6, and 8) was closely affiliated with Thiovulum sp., contained within the ε-Proteobacteria and accounted for ~61% of the SSU rRNA bacterial clone library from Pele's Vents. A second, smaller cluster of PVB OTUs (1 and 11) was closely affiliated with Xanthomonas sp., contained within the γ-Proteobacteria, and accounted for a total of ~27% of the bacterial clone library. The remaining five PVB OTUs each accounted for ~2% of the clones recovered and were affiliated with the following phylogenetic groups: PVB OTU 5 was a member of the Alteromonas group; PVB OTU 12 was a member of the Colwellia assemblage; PVB OTU 4 was loosely determined to be a member of the Thiothrix group, with the endosymbiotic bacteria from Bathymodiolum thermophilus and Calyptogena.
magnifica as nearest relatives; PVB OTU 10B was a member of the Myxobacteria group; and PVB OTU 9A was a member of the Paraphyletic assemblage, with the Octopus Spring microbial mat type K clone as the closest known relative. PVB OTU 7 was determined to be a PCR-generated chimeric structure combined from two described phylotypes detected in this study, thereby decreasing the previously estimated number of major PVB OTUs from 12 to 11.

**INTRODUCTION**

Hydrothermal vent microbial communities are potentially diverse because of a plethora of habitats sustained by both chemical and physical extreme gradients. The most widely accepted (or at least hypothesized) mode of metabolism thought to dominate hydrothermal vent microbial communities is chemolithoautotrophy, principally through the oxidation of reduced sulfur and iron compounds (Jannasch and Mottl, 1985). However, even though numerous metabolic pathways are possible at hydrothermal vent systems, quantitative assessments of these independent metabolic pathways to total community metabolism have not yet been possible (Karl, 1987; Karl, 1995). A recent study of the phylogenetic affiliations of sulfur- and iron-oxidizing bacteria through the use of small-subunit (SSU) rRNA demonstrated the ubiquity of these metabolic pathways throughout the entire domain Bacteria (sensu Woese [Olsen et al., 1994; Woese et al., 1990]), which suggests an early evolutionary development with respect to life on this planet (Lane et al., 1992). These data were also consistent with the original phylogenetic studies using
5S rRNA (Lane et al., 1985; Stahl et al., 1987), which demonstrated that aerobic sulfur bacteria are phylogenetically and taxonomically very heterogeneous.

The majority of deep-sea bacteria either grow slowly or are dormant as a result of the limitation of bioavailable carbon and energy (Morita, 1988). Ultimately, these primarily heterotrophic bacteria derive their meager existence from recalcitrant organic matter that becomes available only after advection or sedimentation from the euphotic zone. In contrast, the first study to examine the metabolic potential of free-living bacteria from hydrothermal vents revealed the presence of chemolithoautotrophic bacteria utilizing readily available, geothermally reduced sulfur compounds at the Galapagos vents (Karl et al., 1980). Later, three distinct physiological groups of putative sulfur-oxidizing bacteria were isolated and described as various strains of obligately chemolithoautotrophic *Thiomicrospira* spp. or as obligately heterotrophic thiobacillus-like and pseudomonad-like microorganisms (Ruby et al., 1981). Microbial mats resembling those formed by the sulfur-oxidizing bacteria *Beggiatoa* and *Thiothrix* spp. were also observed in samples from these vents (Ruby et al., 1981). These and other sulfur- and iron-oxidizing bacteria can grow in habitats containing oxic-anoxic interfaces. The isolation of an obligately anaerobic *Spirochaeta* sp. from a marine hydrothermal vent (Harwood et al., 1982) demonstrates the potential for anoxia in such habitats. In addition, Karl et al. (1988; 1989) conducted physiological studies and microscopic examinations of microbial mats from the Pele's Vents hydrothermal vent system; they found thermophilic bacterial populations, evidence of iron and sulfur oxidation, and observed the microbial mats to be dominated by iron-depositing sheathed bacteria.
The use of molecular biological techniques, especially those that take advantage of the SSU rRNA molecule, has eliminated the dependence upon isolations of pure cultures as a means of studying the diversity and structure of natural microbial communities (Pace et al., 1986; Ward et al., 1992). However, this type of molecular biological strategy also has potential pitfalls and limitations, which investigators must consider and attempt to control. For example, the in vitro generation of chimeric SSU ribosomal DNA (rDNA) structures during polymerase chain reaction (PCR) amplification procedures has recently been documented using mixtures of pure culture isolates (Liesack et al., 1991) and environmentally derived, uncultivated bacteria (Kopczynski et al., 1994; Moyer et al., 1994). In addition to analysis of SSU rRNA secondary structures, an automated procedure for the detection of chimeric SSU rRNA sequences through similarity analysis (CHECK_CHIMERA) is now available through the Ribosomal Database Project (RDP [Larsen et al., 1993]), and the utility of this program is thoroughly discussed by Kopczynski et al. (1994). With the combined use of these quality control procedures, the identities of possible chimeric sequences can be ascertained, and the exact location of a chimeric splicing site can potentially be detected.

In the present study, the phylogenetic diversity of the bacterial community from Pele's Vents, an active deep-sea hydrothermal vent system, located on the summit of Loihi Seamount, Hawaii was examined. The focus was on the operational taxonomic units (OTUs) previously characterized by a restriction fragment length polymorphism (RFLP) distribution analysis of bacterial SSU rRNA genes. In the initial SSU rRNA molecular biological study of the microbial mats at Pele's Vents, it was determined that the bacterial
community was dominated by two OTUs and that several others were present at reduced
levels of abundance (Moyer et al., 1994). In addition, it was reported that the SSU rDNA
clone library generated to represent the bacterial community at Pele's Vents had been
sampled sufficiently to detect a majority of the phylogenetic diversity present in the native
habitat (Moyer et al., 1994). This chapter reports on the sequence analyses of the
SSU rRNA genes from each of the 11 major bacterial OTUs to (i) determine the genetic
relatedness of the OTUs to one another and (ii) determine each OTU's ancestry (phylogeny)
with a database of known SSU rRNA sequences (Larsen et al., 1993).

MATERIALS AND METHODS

Sample collection, generation of bacterial clone library, and RFLP distribution analysis

The collection of hydrothermal vent microbial mat samples, the construction of the
PCR-amplified bacterial clone library, and the screening of SSU rDNA clones through RFLP
distribution analysis and rDNA fingerprinting are described elsewhere (Moyer et al., 1994).

rDNA sequencing

Representative bacterial SSU rDNA clones were sequenced with an automated DNA
sequencer Model 373A (Applied Biosystems, Foster City, CA.). Sequencing was performed
according to the manufacturer's specifications by using plasmid templates with fluorescently
labeled dideoxy terminators and PCR (Taq polymerase) cycle sequencing. Oligonucleotides
used as primers at various positions internal to the bacterial SSU rRNA gene were the same as those described by Lane (1991) and were synthesized and purified as previously described (Moyer et al., 1994).

**Phylogenetic analysis**

Sequences were manually aligned to a database of previously determined SSU rRNA sequences obtained from the RDP (Larsen et al., 1993). Sequence alignments were based on primary and secondary structural considerations (see Appendix C) and were constructed using the GDE multiple sequence editor distributed through the RDP (Larsen et al., 1993). Sequences also were manually aligned into complete secondary structures and were submitted to the CHECK_CHIMERA program to detect for the presence of possible chimeric artifacts (Kopczynski et al., 1994; Larsen et al., 1993). Phylogenetic analyses were restricted to the comparison of highly to moderately conserved nucleotide positions that were unambiguously alignable in all sequences, corresponding to residues 101 to 183, 220 to 451, 482 to 840, 846 to 1005, 1037 to 1133, and 1141 to 1445 (Escherichia coli numbering system). Initial phylogenetic screening was conducted using the DeSoete algorithm (DeSoete, 1983), which fits distance matrix data to an optimal additive tree. Corrected pairwise distances were computed from percent similarities by the Jukes and Cantor method (Jukes and Cantor, 1969), as modified by G.J. Olsen to accommodate the actual base ratios (Weisburg et al., 1989). Final phylogenetic placement was conducted through maximum likelihood analyses with the fastDNAml program (version 1.0.6c) written
by G.J. Olsen and distributed by the RDP (Larsen et al., 1993). This software was derived from J. Felsenstein's DNaml program (version 3.3; part of the PHYLIP package) and uses the generalized two-parameter model of evolution (Kishino and Hasegawa, 1989; Thorne et al., 1991). Final phylogenetic trees were constructed using jumbled orders for the addition of taxa and allowed for the global swapping of branches. Using these parameters, the search for an optimal tree was repeated until the best log likelihood score was reached in at least three independent searches. Bootstrapping methods were used with each dataset so that node reproducibilities for tree topologies could be estimated (Felsenstein, 1985). Each dataset was bootstrapped 500 times with the jumbled addition of taxa, and the search for an optimal tree was repeated until the best log likelihood score was reached in at least two independent searches.

**Nucleotide sequence accession numbers**

The SSU rRNA sequences representing the OTUs for Pele's Vents Bacteria (PVB OTUs) used in the present analysis have been submitted to GenBank and assigned accession numbers U15100 through U15107 and U15111 through U15118.

**RESULTS**

A total of 48 clones were screened using an RFLP distribution analysis, and 12 PVB OTUs were initially detected, each containing an ~1.5 kb SSU rDNA insert (Moyer
et al., 1994). Three representative clones were sequenced *in toto* from OTUs 1, 2 and 3, which contained 12, 23, and 3 clones, respectively. The only other OTU that contained multiple clones was OTU 6, in which two clones were detected and sequenced *in toto*. The remaining eight OTUs were represented by a single clone, and each was sequenced *in toto*. Upon examination of sequence data through the comparison of SSU rRNA secondary structure models and the use of the CHECK_CHIMERA program, it was determined that three of the eight OTUs represented by a single clone were PCR-generated chimeric structures. Two of the chimeric clones (representing OTUs 9 and 10) were still predominantly made up of novel, phylogenetically contiguous segments of DNA sequence (phylotypes). These clones were defined as OTU 9A, which was represented conservatively by a 1,007 bp segment contiguous from the 5' end of the SSU rRNA gene, and OTU 10B, which was represented conservatively by a 790 bp segment contiguous from the 3' end of the SSU rRNA gene. The sequences of each of the remaining segments from the clones representing OTUs 9 and 10 were identical to sequences from similar positions in OTUs 2 and 1, respectively. The third chimeric OTU detected was OTU 7, which was determined to be constructed of segments found in OTUs 2 and 9A (5' end to position ~295 was OTU 9A, and position ~295 to the 3' end was OTU 2). This result lowered the overall estimate of detected OTUs (phylotypes) from 12 to 11, thus eliminating OTU 7 and the sequence segments from OTUs 9B and 10A from further phylogenetic analysis.

The phylogenetic affiliations of all PVB OTUs are summarized in Table 4.1, together with the relative abundance of each respective OTU. By far the most frequently encountered group of OTUs was the OTU 2 cluster, comprised of OTUs 2, 3, 8, and the
Table 4.1. Phylogenetic affiliations and percent recoveries of bacterial SSU rRNA genes from a microbial mat at Pele’s Vents

<table>
<thead>
<tr>
<th>Phylogenetic affiliation</th>
<th>Sequenced clone no.</th>
<th>GenBank no.</th>
<th>RFLP group</th>
<th>Clones recovered (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>e-Proteobacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiovulum group</td>
<td>PVB 7</td>
<td>U15100</td>
<td>PVB OTU 2 (23)</td>
<td>47.9</td>
</tr>
<tr>
<td></td>
<td>PVB 15</td>
<td>U15101</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVB 63</td>
<td>U15102</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVB 10</td>
<td>U15103</td>
<td>PVB OTU 3 (3)</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>PVB 12</td>
<td>U15104</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVB 55</td>
<td>U15105</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVB 28</td>
<td>U15106</td>
<td>PVB OTU 6 (2)</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>PVB 73</td>
<td>U15106</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVB 32</td>
<td>U15107</td>
<td>PVB OTU 8 (1)</td>
<td>2.1</td>
</tr>
<tr>
<td>γ-Proteobacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xanthomonas group</td>
<td>PVB 3</td>
<td>U15111</td>
<td>PVB OTU 1 (12)</td>
<td>25.0</td>
</tr>
<tr>
<td></td>
<td>PVB 5</td>
<td>U15111</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVB 25</td>
<td>U15112</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVB 47</td>
<td>U15113</td>
<td>PVB OTU 11 (1)</td>
<td>2.1</td>
</tr>
<tr>
<td>Alteromonas group</td>
<td>PVB 18</td>
<td>U15114</td>
<td>PVB OTU 5 (1)</td>
<td>2.1</td>
</tr>
<tr>
<td>Colwellia assemblage</td>
<td>PVB 54</td>
<td>U15115</td>
<td>PVB OTU 12 (1)</td>
<td>2.1</td>
</tr>
<tr>
<td>Thiothrix group</td>
<td>PVB 13</td>
<td>U15116</td>
<td>PVB OTU 4 (1)</td>
<td>2.1</td>
</tr>
<tr>
<td>δ-Proteobacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myxobacteria group</td>
<td>PVB 46</td>
<td>U15117</td>
<td>PVB OTU 10B (1)</td>
<td>2.1</td>
</tr>
<tr>
<td>Paraphyletic assemblage</td>
<td>PVB 37</td>
<td>U15118</td>
<td>PVB OTU 9A (1)</td>
<td>2.1</td>
</tr>
<tr>
<td>Chimera</td>
<td>PVB 30</td>
<td>—</td>
<td>PVB OTU 7 (1)</td>
<td>2.1</td>
</tr>
</tbody>
</table>

*As described by the Ribosomal Database Project, version 4.0.

*As determined by Moyer et al. (1994); in that study OTUs were defined through RFLP distribution analysis. Numbers in parentheses indicate the total number of recovered clones contained within each PVB OTU. The designation of PVB OTUs 9A and 10B refers to contiguous phylotypes as described in Materials and Methods.

†Calculated by dividing the number of group-specific bacterial clones by the total number of clones analyzed (n = 48).

‡Chimeric SSU rRNA clone combined from sequences belonging to PVB OTUs 9A and 2.
more distantly related OTU 6. The OTU 2 cluster was closely affiliated with the Thiovulum
group, which is contained within the \textit{\textbf{e-Proteobacteria}} (Fig. 4.1; Table 4.1). The OTU 2
cluster accounted overall for \(\sim 61\%\) of the bacterial SSU rRNA clones recovered from Pele's
Vents. The close association between the OTU 2 cluster and \textit{Thiovulum} sp. was also
confirmed through SSU rRNA secondary structure analysis (data not shown), by which it
was determined that both taxa have the characteristic short-short helical pattern of the
SSU rRNA cruciform structure described by Lane \textit{et al.} (1992). OTUs 2, 3, and 8 are
closely related members of the OTU 2 cluster, which together incurred a total of 18 variable
positions. Of these, two sites exhibited compensatory base changes, where complimentary
bases changed on both sides of a stem structure, thereby preserving the secondary structure
of the SSU rRNA molecule. Perhaps in relation to the high abundance of clones from the
OTU 2 cluster, some of the variability in OTUs 2, 3, and 8 may be accounted for by
microheterogeneity between different \textit{rrn} operons within the same bacterial population. This
explanation is certainly not the case with OTU 6, which was by far the most phylogenetically
distant related member of the OTU 2 cluster (Fig. 4.1). OTU 6 had numerous variable
positions (relative to OTUs 2, 3, and 8), including several compensatory base changes, while
still maintaining the same overall secondary structure characteristic of the Thiovulum group
(data not shown). In addition, the occurrence of OTU 6 was detected twice, and in both
instances the sequence data each clone were identical, as indicated by the common GenBank
accession numbers (Table 4.1).

The second most abundantly represented phylotype was the \textit{PVB OTU 1 cluster},
which was comprised of OTUs 1 and 11, and accounted for \(\sim 27\%\) of the bacterial clone
Figure 4.1. Phylogenetic tree demonstrating relationships of the PVB OTU 2 cluster with \( \varepsilon \)-Proteobacteria as determined by maximum likelihood analysis of SSU rDNA sequences. Numbers at nodes represent bootstrap values (percent) for that node (based on 500 bootstrap resamplings). Outgroups are represented by \( E. \) coli and \( D. \) acetivorans (Rainey et al., 1993) SSU rDNA sequence data. Sequences not determined in this study were provided by the Ribosomal Database Project (Larsen et al., 1993), except as noted. The scale bar represents 0.10 fixed mutations per nucleotide position. Bootstrap values are shown for frequencies at or above a threshold of 50%. 
library recovered from Pele's Vents. The OTU 1 cluster was affiliated with the Xanthomonas group and, in addition, was closely related to several *Thiobacillus* spp., some of which are contained in the Chromatium assemblage and all of which are contained phylogenetically near the root of the γ-Proteobacteria (Fig 4.2; Table 4.1). Sequences from the clones representing OTUs 1 and 11 varied by a single *RsaI* restriction site and had a total of seven variable positions. It is assumed that this level of variation approximates the lower limit of resolution for the RFLP distribution analysis used to determine the OTUs (Moyer *et al.*, 1994). However, the possibility of microheterogeneity between different *rrn* operons within the same bacterial population cannot entirely be ruled out. Again, identical sequences were detected for two of three of the clones constituting PVB OTU 1, as indicated by common GenBank accession numbers (Table 4.1).

Five other phylogenetically distinct PVB OTUs, each represented by a single clone, were also detected, and these cumulatively accounted for ~10% of the SSU rRNA clone library (Table 4.1). Four of these were contained within the *Proteobacteria*, with the first being OTU 5, a member of the heterogeneous Alteromonas group and having *Alteromonas haloplanktis* as the closest described relative (Fig. 4.2). The second was OTU 12, a member of the Colwellia assemblage, which had the environmentally-isolated marine aggregate clone 53 (DeLong *et al.*, 1993) and *Colwellia psychroerythrus* as the closest described relatives (Fig. 4.2). The final OTU affiliated with the γ-Proteobacteria was OTU 4, which was most closely related to the chemolithoautotrophic bacterial endosymbionts from *Bathymodiolus thermophilus* and *Calyptogena magnifica* and thereby was determined to be a member of the Thiothrix group. However, OTU 4 was also closely related to
Figure 4.2. Phylogenetic tree demonstrating relationships of the PVB OTU 1 cluster, PVB OTU 5, and PVB OTU 12 with $\gamma$-Proteobacteria as determined by maximum likelihood analysis of SSU rDNA sequences. Numbers at nodes represent bootstrap values (percent) for that node (based on 500 bootstrap resamplings). An outgroup is represented by *Aquifex pyrophilus* SSU rDNA sequence data. Sequences not determined in this study were provided by the Ribosomal Database Project (Larsen et al., 1993). The scale bar represents 0.10 fixed mutations per nucleotide position. Bootstrap values are shown for frequencies at or above a threshold of 50%.
Alteromonas haloplanktis

Alteromonas macleodii

PVB OTU 5

PVB OTU 12

PVB OTU 1

PVB OTU 11

Xanthomonas maltophilia

Xylella fastidiosa

Thiobacillus ferrooxidans M-1

Aquifex pyrophilus
Thiobacillus ferroxidans strain M-1, and so the possibility of placement in the Chromatium assemblage cannot entirely be ruled out (Fig. 4.3). OTU 10B was affiliated with the Myxobacteria group of the δ-Proteobacteria, having Chondromyces spp. and Polyangium spp. as closest relatives (Fig. 4.4). This result was confirmed by a comparison of several distinguishing shared derived characters (synapomorphies) at the secondary structural level, using the myxobacterial sequence signatures determined by Shimkets and Woese (1992). The final phylotype examined was OTU9A, which was placed in the Paraphyletic assemblage (Fig. 4.5). The Paraphyletic assemblage is deeply rooted in the domain Bacteria (Olsen et al., 1994; Woese et al., 1990). The closest relative to OTU9A was another environmental clone isolate, Octopus Spring microbial mat type K (Weller et al., 1992).

DISCUSSION

At least four generic microbially-based communities associated with hydrothermal vent habitats are known to exist; these include (i) free-living bacterial populations associated with the discharged vent fluids and presumably growing and reproducing within the subseabed strata, (ii) free-living microbial mats growing on surface strata that are exposed to flowing vent waters, (iii) endo- and exosymbiotic associations of microorganisms and vent fauna, and (iv) microorganisms within the deep-sea hydrothermal vent plumes (Karl, 1987; Karl, 1995). In addition, cold seawater surrounds and permeates the entire hydrothermal vent ecosystem and provides physical, chemical, and biological inputs, thereby affecting all of the habitats contained therein. Each of these factors must be considered when one is
Figure 4.3. Phylogenetic tree demonstrating relationships of PVB OTU 4 with \textit{\gamma-Proteobacteria} as determined by maximum likelihood analysis of SSU rDNA sequences. Numbers at nodes represent bootstrap values (percent) for that node (based on 500 bootstrap resamplings). An outgroup is represented by \textit{Aquifex pyrophilus} SSU rDNA sequence data. Sequences not determined in this study were provided by the Ribosomal Database Project (Larsen et al., 1993). The scale bar represents 0.10 fixed mutations per nucleotide position. Bootstrap values are shown for frequencies at or above a threshold of 50%.
Aquifex pyrophilus

Thiobacillus caldus

Thiobacillus tepidarius

hydrothermal isolate NF-13

Thiomicropira sp. L-12

Thiomicropira thyasiris

Chromatium vinosum

Ectothiorhodospira halochloris

Riftia pachyptila symbiont

Thiobacillus ferrooxidans M-1

PVB OTU 4

Bathymodiolus thermophilus symbiont

Calyptogena magnifica symbiont

Thiothrix nivea

Piscirickettsia salmonis

Escherichia coli

Escherichia coli
Figure 4.4. Phylogenetic tree demonstrating relationships of PVB OTU 10B with other myxobacteria as determined by maximum likelihood analysis of SSU rDNA sequences. Numbers at nodes represent bootstrap values (percent) for that node (based on 500 bootstrap resamplings). Outgroups are represented by *E. coli*, *Desulfovibrio desulfuricans*, and *Desulfurella acetivorans* (Rainey et al., 1993) SSU rDNA sequence data. Sequences not determined in this study were provided by the Ribosomal Database Project (Larsen et al., 1993), except as noted. The scale bar represents 0.10 fixed mutations per nucleotide position. Bootstrap values are shown for frequencies at or above a threshold of 50%.
Figure 4.5. Phylogenetic tree demonstrating relationships of PVB OTU 9A with the Paraphyletic assemblage and the Leptospirillum group as determined by maximum likelihood analysis of SSU rDNA sequences. Numbers at nodes represent bootstrap values (percent) for that node (based on 500 bootstrap resamplings). An outgroup is represented by *Aquifex pyrophilus* SSU rDNA sequence data. Sequences not determined in this study were provided by the Ribosomal Database Project (Larsen et al., 1993). The scale bar represents 0.10 fixed mutations per nucleotide position. Bootstrap values are shown for frequencies at or above a threshold of 50%.
Methanobacterium bavaricum

Octopus Spring isolate OPI-2

Thermodesulfovibrio yellowstonii

Octopus Spring microbial mat type K

PVB OTU 9A

Octopus Spring locus I-25

Octopus Spring microbial mat type L

Synergistes jonesii

Aquifex pyrophilus
designing an appropriate sampling scheme that will avoid sampling bias, especially in view of the constraints imposed when using a submersible for sample collection.

Phylogenetic analyses of SSU rDNA clones provide a method for assessing the structure and diversity of a microbial community without introducing the bias inherent in pure culture isolation techniques. However, as with any sampling-detection procedure, the potential for introducing some bias and selectivity still remains. When molecular biological strategies of this type are used, there are potential pitfalls and limitations in addition to the previously described problem of PCR-generated chimeric SSU rDNA structures. Another potential limitation can occur when the efficient and unbiased extraction of nucleic acids from natural microbial communities is attempted. The presence of large concentrations of exopolymeric substances and mineral deposits in hydrothermal vent sample materials may reduce the extraction efficiency or bias the yields of total nucleic acids from natural bacterial populations (Karl, 1995). This problem can potentially be circumvented through the combined use of French pressure and enzymatic cellular lysis, microscopic examination of sample slurry, and purification steps to maximize the recovery of total nucleic acids (Moyer et al., 1994). The differential PCR-mediated amplification and blunt-end cloning of SSU rRNA genes are additional sources of potential selective bias. The use of multiple PCRs and ligation reactions in the generation of a clone library may act to diminish these effects, although the complete elimination of these potential biases is not yet possible and thus continues to be a potential disadvantage of these cloning methods (Moyer et al., 1994).

The wide diversity among the bacterial phylotypes recovered in this study suggests that there was no unique bias during the extraction, amplification, and cloning procedures.
Chimeric structures accounted for only ~6% (n=3) of the total number of clones from the bacterial SSU rDNA clone library, and the points of recombination within each of the chimeric SSU rDNA clones were determined. Two of the chimeric clones were primarily comprised of novel phylotypes (OTUs 9A and 10B), and only a single clone (OTU 7) was recombined completely from other described phylotypes from this study. Additionally, no clones detected more than once through the RFLP distribution analysis were determined to contain chimeric structures. As described above, molecular biological techniques may involve biases that are not yet fully understood, but these biases are much less encumbering than those imposed by the requirements of selective enrichment and pure culture isolation. These and other data (DeLong et al., 1993; Fuhrman et al., 1993; Giovannoni et al., 1990; Schmidt et al., 1991; Ward et al., 1990) demonstrate the effectiveness of evaluating the structure and diversity of a microbial community from environmentally derived mixed populations using SSU rRNA-based molecular biological techniques.

This study assessed the bacterial community structure and diversity at Pele's Vents through the phylogenetic analysis of bacterial OTUs (Table 4.1). Assuming that the SSU rDNA clone library approximated the distribution of populations from the bacterial mat community, then the mats at Pele's Vents are dominated (~61%) by the populations of bacteria represented in the PVB OTU 2 cluster. The OTU 2 cluster is most closely related to *Thiovulum* sp. (Fig. 4.1), a mesophilic obligate chemolithoautotrophic member of the *e-Proteobacteria*, and may therefore represent a source of autotrophic productivity for the microbial community at Pele's Vents. *Thiovulum* sp. has been studied extensively in the laboratory with enrichment cultures (La Rivière et al., 1991; Wirsen and Jannasch, 1978).
The most characteristic features of *Thiovulum* sp. is that it grows in veils and webs (held together by a polysaccharide matrix) and that it is found in sharply localized white masses in habitats that encompass the interface between sulfide and oxygen (La Rivière et al., 1991). Moreover, the suggestion that the OTU 2 cluster represents bacterial populations like *Thiovulum* sp. is supported not only phylogenetically but also by the observation of large clumps of white "streamers" of bacteria surrounding the orifices of Pele's Vents. The veils or webs formed by *Thiovulum* sp. form a stabilized microenvironment by creating an unstirred boundary layer, which minimizes the mixing of sulfide and oxygen and thus also minimizes the chemical oxidation of sulfide (Jørgensen and Revsbech, 1983). The venting waters at Pele's Vents contain extremely high concentrations of total dissolved CO$_2$ (ca. 300mM), which are more than 100 times greater than the concentration at the Galapagos Rift Vents (Edmond et al., 1987). Enriched cultures of *Thiovulum* sp. demonstrated significant sulfide-stimulated assimilation of dissolved CO$_2$ in addition to not assimilating Casamino Acids, acetate, glutamate, or mannitol, further suggesting a chemolithoautotrophic mode of metabolism (Wirsen and Jannasch, 1978). The present study is the first report of *Thiovulum*-like bacterial populations from a deep-sea hydrothermal vent habitat. Since *Thiovulum* sp. has yet to be grown in pure culture, representative isolates from Pele's Vents are not expected; however, future studies using enrichment culture and *in situ* oligonucleotide hybridization techniques may provide additional physiological and ecological information.

The second most highly abundant group (~27%) was represented by the PVB OTU 1 cluster, which is a member of the Xanthomonas group (as defined by the RDP)
but is also closely related to several *Thiobacillus* spp., some of which are contained in the Chromatium assemblage (Fig. 4.2); all of these taxa are contained in the \(\gamma\)-Proteobacteria. The closely related *Thiobacillus* spp., all of which are mesophilic obligate chemolithoautotrophs, include *Thiobacillus hydrothermalis*, isolated from a deep-sea hydrothermal vent in the Fiji Basin (Durand *et al.*, 1993). Xanthomonads are known for their production of characteristic yellow pigments (xanthomonadins) and their production of copious amounts of complex extracellular polysaccharides (Palleroni, 1991). This phylotype (the PVB OTU 1 cluster) is considered to represent a free-living addition to the pseudomonad rRNA similarity group V, as described by Palleroni *et al.* (1973). This phylotype may contribute to the observed yellow color of the mats at Pele's Vents, as well as to the complex extracellular polysaccharides found therein. However, it must be realized that production of yellow pigments and complex extracellular polysaccharides are characteristics which are phylogenetically widespread and which are plastic even among individual bacterial strains.

Two additional, phylogenetically distinct PVB OTUs, each represented by a single clone, were contained in the taxonomically defined, highly heterogenous genus *Alteromonas*, which falls within the \(\gamma\)-Proteobacteria and is composed primarily of heterotrophic marine bacteria. The first phylotype detected was OTU 5, a member of the *Alteromonas* group (as defined by the RDP), which had *Alteromonas haloplanktis* and an as-yet-undescribed purple bacterium as the closest relatives (Fig. 4.2). The entire *Alteromonas* group taxonomically centers around *Alteromonas haloplanktis* and includes most of the recognized *Alteromonas* species (Gauthier and Breittmayer, 1991). The second
the phylotype detected was OTU 12, a member of the Colwellia assemblage (as defined by the RDP), which had the environmentally isolated marine aggregate clone 53 (DeLong et al., 1993) and *Colwellia psychoerythrus* as the closest described relatives (Fig. 4.2). This group taxonomically centers around *Alteromonas macleodii* and is clearly separate from the other *Alteromonas* species. Most *Alteromonas* species have been isolated from seawater, including many from the open ocean around the Hawaiian archipelago (Baumann et al., 1972). However, no phylotypes from the cyanobacteria, prochlorophytes, or *α*-Proteobacteria (e.g. the SAR 11 cluster), which are known to occur in high abundance within the surface mixed layer of the Atlantic and Pacific Oceans were detected (DeLong et al., 1993; Fuhrman et al., 1993; Giovannoni et al., 1990; Schmidt et al., 1991). It is concluded from these observations that OTUs 5 and 12 are genuine members of the bacterial mat community contained within the hydrothermal ecosystem at Pele's Vents. In addition, it is hypothesized that hydrothermal vent habitats, in general, may be an important source of carbon and energy for many of these types of heterotrophic marine bacteria, which are otherwise in the starvation-survival state when found throughout the water column (Morita, 1982; Morita, 1985). An alternative hypothesis (the two are not mutually exclusive) is that these particular bacteria may reside in the microbial mats as a result of the cold seawater circulation through this hydrothermal vent system. This hypothesis is supported by the ecological distribution of planktonically isolated psychrophiles and barophiles (Morita, 1975), the closest known relatives of OTUs 5 and 12, which are contained almost exclusively in the genera *Colwellia* and *Alteromonas* (Deming et al., 1988; Gauthier and Breitmayer, 1991).
The final phylotype affiliated with the $\gamma$-Proteobacteria was PVB OTU 4 (Fig. 4.3). This phylotype was most closely related to the chemolithoautotrophic bacterial endosymbionts from Bathymodiolus thermophilus and Calyptogena magnifica and thereby was determined to be a member of the Thiothrix group (as defined by the RDP). However, the possibility that OTU 4 belongs to the Chromatium assemblage (as defined by the RDP), could not be ruled out because of the similarity to Thiobacillus ferrooxidans strain M-1. Gill endosymbionts from lucinid clams and the trophosome endosymbiont from Riftia pachyptila have a lineage phylogenetically distinct from the one associated with the gill endosymbionts from the hydrothermal vent bivalves Bathymodiolus thermophilus and Calyptogena magnifica (Distel et al., 1988). These two endosymbiont lineages are contained in the Chromatium assemblage and Thiothrix group, respectively. Thus, the phylogenetic uncertainty about OTU 4 indicates that this phylotype either belongs to one of these lineages or may be an ancestral phylotype of both lineages. Pele's Vents do not have the luxuriant macrofauna found at other hydrothermal vent systems; however, recently a novel species of bresiliid shrimp (Williams and Dobbs, 1995) and a novel pogonophoran worm (pers. comm. Dr. Robert Vrijenhoek) have been detected in association with this hydrothermal vent system. At this time, it is not known whether the OTU 4 phylotype is an endo- or exosymbiotic partner with either of these species of macrofauna.

The PVB OTU 10B phylotype was closely affiliated with the myxobacteria, which are contained in the $\delta$-Proteobacteria, and had Chondromyces spp. and Polyangium spp. as the closest relatives (Fig. 4.4). This result is unusual since no myxobacteria are known to grow at the salt concentrations found in seawater (Reichenbach and Dworkin, 1991),
even though several species of myxobacteria have been isolated from seashore sediments (Brockman, 1967; Ruckert, 1975). The principal habitats for myxobacteria are associated with soils and are ubiquitously distributed worldwide. Almost all myxobacteria are "micropredators" (Singh, 1947) and are attracted to habitats with rich microbial communities (Reichenbach and Dworkin, 1991). If the OTU 10B phylotype is indeed a typical myxobacteria-like organism, then the relatively concentrated biomass within the microbial mats at Pele's Vents may provide a niche for such a micropredator, thereby constituting the next step in trophic level structure.

The PVB OTU9A phylotype was included in the Paraphyletic assemblage, which is itself deeply rooted in the domain Bacteria (Fig. 4.5). The closest relative to OTU9A was another environmental clone isolate, Octopus Spring microbial mat type K (Weller et al., 1992). This result was most intriguing because this Octopus Spring clone (type K), which occurred only once in a bacterial community consisting of 15 detected phylotypes, came from a predominantly cyanobacteria-based (i.e. photoautotrophic) rather than a Thiovulum-based (i.e. chemoautotrophic) community (Ward et al., 1992). The inability to relate these phylotypes with others from the domain Bacteria suggests that these microbial mat communities are inhabited by additional as-yet-undiscovered phylotypes (i.e., without culturable analogs). Ward et al. (1992), speculated that these phylotypes (based on their overall low abundance) may represent community members that occupy a secondary trophic level; diversity within the community may increase as does the variety of substrates made available by primary producers to microorganisms occupying higher trophic levels. However, the present inability to relate these phylotypes to better-defined phylogenetic or
physiological groups makes it difficult to determine the ecological significance of these results.

This study has focused on the phylogenetic analyses of SSU rRNA sequence data from previously defined OTUs (Moyer et al., 1994) to estimate the genetic diversity contained within the bacterial community from Pele's Vents. This approach is free of the bias introduced by classical microbial cultivation techniques but may still contain some selective biases from the types of molecular biological techniques as used herein. It was determined that the community is primarily composed of a cluster of Thiovulum-like phylotypes and secondarily composed of a cluster of xanthomonad-like phylotypes, assuming the nearly equivalent recovery of the bacterial community members. With this assumption in mind, it is interesting to note that both of the dominant PVB OTU clusters (1 and 2) contain a single dominant phylotype. The remaining phylotypes represent a wide variety of SSU rRNA genetic diversity spanning the domain Bacteria. The phylogenetic analysis of recovered SSU rDNA clones from the domain Archaea are reported elsewhere (Moyer et al., 1995). Further examination of the bacterial community from Pele's Vents using in situ hybridization techniques will enable the confirmation and localization of specific phylotypes and begin to delineate their spatial and temporal variability. Pele's Vents is the first hydrothermal vent microbial community to be dissected with this type of molecular biological approach, making it a model against which to compare other microbially based hydrothermal vent communities.
REFERENCES


CHAPTER 5

MOLECULAR PHYLOGENY OF DEEP-SEA HYDROTHERMAL VENT ARCHAEA

ABSTRACT

Through the examination of SSU rRNA genes, the molecular phylogeny of the domain *Archaea* (one of the three major lineages of life) was analyzed from a microbial mat community at an active, deep-sea hydrothermal vent ecosystem located at Pele's Vents on the summit of Loihi Seamount, Hawaii. The genes were isolated from purified microbial mat DNA by PCR, blunt-end cloned into a plasmid vector, and sequenced. The derived archaeal sequences were then used to infer the evolutionary relationships between these microbial mat community members and their closest known relatives. A cluster of three phylogenetically similar PVA (Pele's Vents *Archaea*) clones were detected, which were contained in the archaeal Group I lineage of the marine *Crenarchaeota*. A single PVA clone was contained in the archaeal Group II lineage of the marine *Euryarchaeota*. All four of the PVA clones appear to be novel and constitute the discovery of new archaeal taxa. On the basis of these results, it is suggested that hydrothermal ecosystems, such as Pele's Vents, may act as a source of marine archaeoplankton (a recently discovered component of marine picoplankton) detected from oceans around the world.
INTRODUCTION

Over the past three years there has been an increase in the discovery of archaeal diversity from aquatic habitats using molecular biological techniques focusing on SSU rRNA (small-subunit ribosomal RNA) phylogeny. The first discoveries came from coastal marine habitats from both the Pacific and Atlantic Oceans and two novel lineages were detected as represented by the marine archaeal Groups I and II (DeLong, 1992; Fuhrman et al., 1992). The Group I lineage is contained within the Crenarchaeota kingdom and the Group II lineage is contained within the Euryarchaeota kingdom. Both of these marine archaeal lineages are deeply-rooted within the Archaea domain sensu Woese et al. (1990) and act to unify the previously defined Crenarchaeota and Euryarchaeota lineages. Additional increases in archaeal diversity have also been reported from hot spring habitats at Yellowstone National Park (primarily for the Crenarchaeota; Barns et al., 1994) and from a deep-sea holothurian (marine archaeal Group I; McInerney et al., 1995). Incredibly, DeLong et al. (1994) have demonstrated that as much as 30% of the surface marine picoplankton from Antarctic coastal waters can be accounted for by archaeal members of the community (archaeoplankton) and that the majority of these (9 of 14 SSU rDNA clones) were associated with the marine archaeal Group I lineage, while the remaining clones were associated with the marine archaeal Group II lineage. This study reports the discovery of marine archaeal community members found in microbial mats from the deep-sea hydrothermal vent ecosystem located at Pele’s Vents on the summit of the mid-plate hotspot volcano, Loihi Seamount, Hawaii.
All isolates from hyperthermophilic deep-sea hydrothermal vents belong to the domain *Archaea*, with the exception of isolates from the *Thermotogales* and their closest relatives, which belong deeply-rooted in the domain *Bacteria*. Archaeal enrichment culture isolations have primarily been achieved with samples taken from the hydrothermal vent habitats located at the East Pacific Rise (21°N and 11°N), Juan de Fuca Ridge, and Guaymas Basin (Karl, 1995). So far, all archaeal isolates have demonstrated the “classic” hyperthermophile phenotype (e.g., thriving at temperatures between 80 and 110°C and are unable to grow below 60°C). In addition, Huber et al. (1990) used DNA/DNA hybridization experiments with enrichment cultures generated from hydrothermal plume samples taken during an eruption event at Macdonald Seamount (another mid-plate hotspot volcano from the South Pacific) to demonstrate the occurrence of members of the hyperthermophilic archaeal genera, *Pyrodictium*, *Pyrococcus*, *Archaeoglobus*, and *Thermococcus*. In all of these enrichment culture studies, representatives from the (presumably mesophilic) marine archaeal groups I or II have yet to be detected or isolated. No known hyperthermophile has yet to be found within either of the marine archaeal group I or II lineages.

Due to the resistance to cultivation of the majority of microorganisms found in the environment, molecular biological techniques have provided an alternative approach to the assessment of the diversity of naturally occurring microbial communities. By taking advantage of the SSU rRNA molecule, a sequence-based phylogenetic approach has been used extensively to detect the presence of previously unknown microorganisms (Pace et al., 1986; Ward et al., 1992). This approach has been especially important for the marine
archaeal Group I and II lineages, as no culturable analogs have yet to be identified from any marine habitat (i.e., planktonic or hydrothermal). Comparative phylogenetic analyses of these recovered sequences can then be used to determine the evolutionary relationships between members of the community and cultivated microbial taxa. These results allow inferences to be made for otherwise unknown microorganisms, based on the properties of their closest known affiliated relatives. In addition, these novel sequences can then be further used to design specific oligonucleotide probes to SSU rRNAs to facilitate the quantitative determination or estimated abundance of that population within a given community.

**MATERIALS AND METHODS**

*Sample collection and generation of archaeal clone library*

The collection of hydrothermal vent microbial mat samples by submersible, construction of the PCR-amplified archaeal clone library, and the screening of SSU rDNA clones are described elsewhere (Moyer et al., 1994; Moyer et al., 1995). PCR primers for the initial amplification of archaeal SSU rDNA templates were analogous to those used by DeLong (1992); priming sites were identical. The forward primer was synthesized as follows: \[5'-\text{TTCYGGTGTACCGCCIGA}\], where Y = C and T, and I = inosine. The reverse primer was synthesized as follows: \[5'-\text{YCCGGCGTGTACCGCCIGA}\], where Y = C and T, and (I/C) = inosine and C at a single position. In addition, the 5' ends of both
primers were constructed with Phosphalink Amidite (Applied Biosystems, Foster City, CA), which aids in the blunt-end cloning process (i.e., synthetic phosphorylation). Oligonucleotides were synthesized and purified as previously described (Moyer et al., 1994).

*rDNA sequencing*

Representative archaeal SSU rDNA clones were sequenced using an automated DNA sequencer, Model 373A (Applied Biosystems, Foster City, CA.). Sequencing was performed according to the manufacturer's specifications using plasmid templates with fluorescently labeled dideoxy-terminators and PCR (Taq polymerase) cycle sequencing. Oligonucleotides used as primers at the various positions internal to the archaeal SSU rRNA gene were the same as those described by Lane (1991), with the exception of the universal priming site located between positions 519 and 536 (Escherichia coli reference numbers), which required redesigning (i.e., slightly increased degeneracy) to accommodate the PVA OTU 2 cluster sequences. The corresponding primers were designed to accommodate this otherwise universally conserved (i.e., similar across all forms of life) site as follows: 519r [5'-G(I/C)TTACCGCGGC(GCTG] and 536f [5'-CAGCMGCCGCGGTA(A/I)IC], where M = A and C, I = inosine, and (I/C) = inosine and C at a single position. Oligonucleotides were synthesized with an automated DNA synthesizer and thin-layer chromatography purified as previously described (Moyer et al., 1994).

117
**Phylogenetic analysis**

Sequences were manually aligned to a database of previously determined SSU rRNA sequences obtained from the ribosomal database project (RDP [Maidak *et al.*, 1994]) or provided by E.F. DeLong in the case of the Group I and II marine archaeal sequences. Sequence alignments were based on primary and secondary structural considerations (see Appendix C) and were constructed using the GDE multiple sequence editor distributed through the RDP (Maidak *et al.*, 1994). Sequences were also manually aligned into complete secondary structures as well as submitted to the CHECK_CHIMERA program to detect for the presence of possible chimeric artifacts (Kopczynski *et al.*, 1994; Moyer *et al.*, 1995). Phylogenetic analyses were restricted to the comparison of highly to moderately conserved nucleotide positions that were unambiguously alignable in all sequences. Phylogenetic analysis was conducted using the DeSoete algorithm (DeSoete, 1983), which fits distance matrix data to an optimal additive tree. Corrected pairwise distances were computed from percent similarities by the Jukes and Cantor method (Jukes and Cantor, 1969), as modified by G.J. Olsen to accommodate the actual base ratios (Olsen, 1987; Weisburg *et al.*, 1989).

**Nucleotide sequence accession numbers**

The SSU rDNA sequences representing the OTUs for Pele's Vents *Archaea* (PVA OTUs) used in the present analysis will be submitted to GenBank for the assignment of accession numbers prior to submission for publication.
RESULTS

Through a cursory examination of an archaeal specific SSU rDNA clone library generated from samples taken at Pele's Vents, four clones were selected after successful primary restriction with *BamHI* and *PstI* as described by Moyer *et al.* (1994). These four clones were then sequenced *in toto*, for a total of ~915 bp each. Upon examination of the entire cloned rDNA sequence through the comparison of SSU rRNA secondary structure models and the use of the CHECK_CHIMERA program, it was determined that each of these clones was a phylogenetically contiguous DNA sequence (phylotype). These tests indicate that all sequences analyzed were free of any chimeric artifacts (data not shown).

The phylogenetic affiliations of the four PVA OTUs (Pele's Vents *Archaea* Operational Taxonomic Units) are described in Fig. 5.1. These archaeal clones have only other environmentally isolate clones as nearest relatives from the following locations: Antarctica, Santa Barbara Channel, Woods Hole, and the Oregon Coast. A cluster of three phylogenetically similar PVA OTUs were detected (i.e., PVA OTU 2 cluster), which were contained in the archaeal Group I lineage of the marine *Crenarchaeota*. A single PVA OTU was contained in the archaeal Group II lineage of the marine *Euryarchaeota*. Although the four PVA OTUs were contained within two discrete phylogenetic lineages, they each represent the detection of novel archaeal taxa. Through the detection and examination of archaeal SSU rRNA clones, it is assumed that these archaeal clones represent additional components contained within the community structure of the microbial mats at Pele's Vents. Currently, none of these archaeal lineages have any known culturable analogs as phylogenetically affiliated relatives.
Figure 5.1. Phylogenetic tree demonstrating relationships of the PYA OTU 2 cluster with the marine archaeal Group I and PVA OTU 1 with the marine archaeal Group II as determined by the DeSoete algorithm (DeSoete, 1983) which fits distance matrix data to an optimal additive tree. This is an unrooted tree, with outgroups represented by *Aquifex pyrophilus* and PVB OTU 9A SSU rDNA sequence data. Sequences not determined in this study were provided by Edward DeLong (Group I and II marine *Archaea*) or the Ribosomal Database Project (Maidak et al., 1994). The scale bar represents 5.00 fixed mutations per 100 nucleotide positions.
DISCUSSION

The genetic diversity of archaeal SSU rDNA clones recovered from the microbial mats at Pele's Vents is limited, in striking contrast to the broad phylogenetic diversity found in bacterial SSU rDNA library originating from these samples (Moyer *et al.*, 1995). This overall pattern is similar to that observed in pelagic marine habitats from the world's oceans, where currently only two lineages of marine archaea (Groups I and II) have been recovered (DeLong, 1992; DeLong *et al.*, 1994; Fuhrman *et al.*, 1992), compared to the numerous lineages of pelagic marine bacteria (Fuhrman *et al.*, 1992; Giovannoni *et al.*, 1990; Schmidt *et al.*, 1991). Classically isolated *Archaea* (i.e., culturable isolates) have been divided into two kingdoms by SSU rRNA phylogenetic analyses (Woese and Olsen, 1986; Woese, 1991): the *Crenarchaeota*, which contains a single phenotype of the sulfur-metabolizing extreme thermophiles and the *Euryarchaeota*, which contains four phenotypes of extreme halophiles, methanogens, thermophilic sulfate reducers, and additional sulfur-metabolizing extreme thermophiles. Each marine archaeal lineage is distinct from any of the previously cultured archaeal group, although both the *Crenarchaeota* and the *Euryarchaeota* kingdoms contain one of these marine archaeal lineages. All methanogens and most extreme thermophiles are strict anaerobes, in contrast to most of the oxygenated marine environment. It may be possible that members of these archaeal lineages grow anaerobically in marine habitats such as hydrothermal vents, submarine hydrocarbon seeps, inside marine snow particles, or within marine sediments (specialized niches devoid of oxygen). Of these potential anaerobic habitats, hydrothermal vents and submarine hydrocarbon seeps are
replete with bioavailable energy sources. However, it may also be that these archaeal lineages are facultative anaerobes or even aerobes (with respect to their physiology), due to their high abundance in oxygenated seawater. It is however, unlikely that either of these marine archaeal lineages are hyperthermophiles because of their relatively low SSU rDNA G + C ratios (DeLong, 1992), and because of the absence of high temperature (>100°C) effluent waters at the Pele’s Vents habitat (see Appendix A for temperature data record from Pele’s Vents).

At least four generic microbially-based communities associated with hydrothermal vent habitats are known to exist; these include (i) free-living bacterial populations associated with the discharged vent fluids and presumably growing and reproducing within the sub-seabed strata, (ii) free-living microbial mats growing on surface strata that are exposed to flowing vent waters, (iii) endo- and exosymbiotic associations of microorganisms and vent fauna, and (iv) microorganisms within the deep-sea hydrothermal vent plumes (Karl, 1987; Karl, 1995). In addition, cold seawater surrounds and permeates the entire hydrothermal vent ecosystem and provides physical, chemical, and biological inputs thereby affecting all of the habitats contained therein. During the studies of the bacteria at Pele’s Vents (Moyer et al., 1994; Moyer et al., 1995), it was hypothesized that hydrothermal vent habitats, in general, may be an important source of bioavailable carbon and energy for many of these free-living types of marine bacteria, which are otherwise in the starvation-survival state (i.e., viable but nonculturable) when found throughout the water column. An alternative hypothesis, which is not mutually exclusive, is that these free-living bacteria may reside in the microbial mats as a result of the cold seawater circulation through this hydrothermal vent.
This hypothesis is supported by the ecological distribution of planktonically isolated psychrophiles and barophiles (Morita, 1975). These represent the closest relatives of PVB OTUs 5 and 12, which were bacterial clones also detected from the microbial mats at Pele’s Vents (Moyer et al., 1995).

These results suggest that hydrothermal ecosystems, such as Pele’s Vents, may act as a potential source of marine archaeoplankton (a recently discovered component of marine picoplankton) detected from oceans around the world. Alternatively, these hydrothermal ecosystems may act as a sink, with deep seawater percolating through the microbial mats and possibly depositing and concentrating archaeoplankton. Although the evidence for Pele’s Vents acting as a source or sink for these archaeal lineages remains to be ascertained, the archaeal SSU rRNA clones detected at this location strongly supports a link between marine hydrothermal and pelagic habitats. Moreover, this study obviates the necessity for future physiological studies of culturable isolates from these phylogenetically distinct archaeal lineages.
REFERENCES


125


CHAPTER 6

CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

SUMMARY

The overall goal of this study was to apply recently developed ribosomal-based molecular biological techniques to address specific questions concerning microbial community structure and diversity from an active deep-sea hydrothermal vent ecosystem. This goal was achieved through the determination and categorization of operational taxonomic units (OTUs) constructed via polymerase chain reaction (PCR) and defined by a restriction fragment length polymorphism (RFLP) distribution analysis. Bacterial and archaeal OTUs were then described based on their respective phylogeny through the analysis of small-subunit ribosomal DNA (SSU rDNA) sequence data. This overall approach provided a culture-independent estimate of the abundance of individuals within an OTU, the number of OTUs present within the community, and the genetic relatedness among OTUs (i.e., community structure and phylogenetic diversity).

In Chapter 2, after determining a need for the further examination of the individual tetrameric restriction enzymes (TREs) used in the generation of RFLP distributions for SSU rDNA clone libraries, the efficacy of 10 TREs was assessed through a computer-simulated analysis across more than 100 bacterial SSU rRNA genes. It was determined that three groups of TREs were present, and that one group (containing the TREs HhaI, RsaI,
and BstUI) was the most efficacious at detecting and separating bacterial SSU rRNA genes (i.e., classifying OTUs) based on their previously determined molecular phylogeny.

In Chapter 3, the construction, screening, and classification of a bacterial SSU rDNA clone library from microbial mat samples taken at the hydrothermal vent system, Pele's Vents, was addressed. The library of 48 clones was restricted with two sets of TREs and a total of 12 OTUs were detected with the rank order abundance of each tabulated. The two most abundant OTUs accounted for ~73% of the total clones examined, and one was approximately twice as abundant as the other. The remaining 10 OTUs accounted for ~27% of the total clones, which were evenly distributed with between one and three clones each. The cumulative OTU distribution for the 48 clones demonstrated that the majority of taxa represented in the clone library were detected, which is also assumed to be an indicator of bacterial diversity (i.e., rarefaction analog). SSU rDNA fingerprinting of individual clones with universal oligonucleotide probes confirmed (i) each clone's proper placement in respective OTUs, and (ii) that the defining RFLP analysis had been correctly conducted on SSU rDNAs in each case.

In Chapter 4, the phylogenetic diversity of the bacterial SSU rRNA genes from previously defined OTUs (hereafter abbreviated PVB OTUs) was examined. A cluster of phylogenetically related PVB OTUs (2, 3, 6, and 8) was closely affiliated with *Thiovulum* sp., contained within the *e*-Proteobacteria and accounted for ~61% of the total SSU rDNA bacterial clone library from Pele's Vents. A second, smaller cluster of PVB OTUs (1 and 11) was closely affiliated with *Xanthomonas* sp., contained within the *γ*-Proteobacteria and accounted for a total of ~27% of the bacterial clone library. The remaining five PVB OTUs
each accounted for ~2% (i.e., detection of a single clone) of the clones recovered and were affiliated with the following phylogenetic groups: PVB OTU 5 was a member of the Alteromonas group; PVB OTU 12 was a member of the Colwellia assemblage; PVB OTU 4 was loosely determined to be a member of the Thiothrix group, with the endosymbiotic bacteria from Bathymodiolus thermophilus and Calyptogena magnifica as nearest relatives; PVB OTU 10B was a member of the Myxobacteria group; and PVB OTU 9A was a member of the Paraphyletic assemblage, with the Octopus Spring microbial mat type K clone as the closest known relative. PVB OTU 7 was determined to be a PCR-generated chimeric structure combined from two described phylotypes detected in this study, thereby decreasing the previously estimated number of major PVB OTUs from 12 to 11.

In Chapter 5, the phylogenetic diversity of the archaeal SSU rRNA genes from an archaeal specific clone library was examined. The molecular phylogeny of the domain Archaea (one of the three major domains of life) is by far the least understood and has received much attention with the advent of the discovery of a high abundance of marine archaeoplankton from around the world’s oceans, most notably the surface waters of Arthur Harbor, Antarctica (up to ~30% of the total picoplankton). A cluster of three phylogenetically similar PVA (Pele’s Vents Archaea) clones was detected, members of which were contained in the archaeal Group I lineage of the marine Crenarchaeota. A single PVA clone was contained in the archaeal Group II lineage of the marine Euryarchaeota. All four of the PVA clones appear to be novel and constitute the discovery of independent archaeal taxa. On the basis of these results, it is hypothesized that hydrothermal ecosystems such as Pele’s Vents may act as a source for marine
archaeoplankton. Alternatively, these hydrothermal ecosystems may be acting as a sink, with deep seawater percolating through the microbial mats and possibly depositing and concentrating archaeoplankton.

This dissertation has successfully combined the fields of oceanography and microbiology, focusing on topics in biodiversity, microbial ecology, evolution, phylogeny, and biotechnology to produce a novel view of microbial community structure and phylogenetic diversity from a hydrothermal vent system. This view has been accomplished through the analysis of SSU rRNA genes, which have been used as descriptors of individual populations and their relative evolutionary histories. This is analogous to the “taking of the census” (i.e., community structure) and the “tracing of the genealogical roots” (i.e., phylogenetic diversity) of the microbial populations contained in the microbial mats at the hydrothermal habitat Pele’s Vents.

FUTURE RESEARCH DIRECTIONS

The next step in addressing community structure issues for the microbial mat from Pele’s Vents and the surrounding hydrothermal habitats located on Loihi Seamount is to directly isolate rRNA for the purpose of oligonucleotide probing by the method suggested by Giovannoni et al. (1990). Once group specific probes are constructed, then direct quantification of taxa-specific targets can be achieved at the RNA level and these results can be compared to those already acquired at the DNA level. The isolation of rRNA has been attempted several times during the current study using methods that focus on both the
physical and the chemical isolation of rRNA from bulk extracted nucleic acids, but with only minor success.

The overall process of rRNA isolation can be broken down into four steps; (i) cell lysis, (ii) inactivation of nucleases, (iii) extraction from environmental matrix, and (iv) purification. The problem to date has been in the final purification process (though the original French pressure cell technique for extraction has the added limitation of being extremely laborious). It has been determined that high concentrations of mucopolysaccharides and iron-rich nontronite clays complex and co-precipitate with the nucleic acids, making the process of nucleic acid purification especially difficult. However, high quality DNA (pure and relatively unnicked) has been isolated repeatedly. The reasons for this result are primarily due to the generally robust nature of DNA relative to RNA and because of the DNA specific purification method using hexadecyltrimethyl ammonium bromide (i.e., CTAB) as described by Moyer et al. (1994).

Currently, I have been working to overcome these limitations by two technique modifications. First, I have successfully adapted the use of glass bead-mill homogenization (Moré et al., 1994) to disrupt the cells of these “hard-to-break” microbial mats (as judged by microscopic examination). This homogenization technique has allowed for the additional advantage of faster sample processing. Second, I am currently testing the use of Sephadex gel as a separation matrix to overcome the rRNA purification difficulty, a technique which has been used successfully with soil and sediment samples (Moran et al., 1993; Tsai and Olson, 1992).

The next phase of examination will focus on the design of oligonucleotide probes
specific to the PVB OTU 2 and 1 cluster phylotypes, respectively. These two lineages, which were novel at the time of detection, have been observed more recently using similar rDNA-based molecular biological techniques to isolate SSU rDNA clones from habitats likely to sustain iron- and sulfur-oxidizing bacteria. First, clones from the PVB OTU 2 lineage (e.g., *Thiovulum*-like) have since been characterized from epibiotic bacteria associated with the hydrothermal vent polychaete, *Alvinella pompejana*, from the Elsa vent site (13°N) along the East Pacific Rise (Haddad et al., 1995). Second, clones from the PVB OTU 1 lineage (e.g., *Xanthomonas*-like) have since been detected in microbial populations associated with a deep anaerobic alkaline aquifer in southeastern Washington state (Department of Energy’s Hanford site), which has also been determined to be dominated by a sulfidogenic microbial metabolism (Fry et al., 1995; Ogram et al., 1995). These results demonstrate that the novel bacterial lineages that I have discovered are not limited to the Pele’s Vents habitat, and may have world-wide biogeochemical as well as ecological implications.

Finally, my overall long-range goal is to compare and contrast the microbial community structure and diversity of other hydrothermal vent systems with that of Pele’s Vents, using the latter as a first-order model. As a starting point, I would like to examine the relationships between the microbial communities from the geologically contrasting settings of hotspot and ridge-axis derived hydrothermal vent systems. This work would include a comparison of mesophilic and hyperthermophilic microbial communities. In addition, I would like to characterize the spatial and temporal variability associated with the ephemeral nature of hydrothermal vent habitats and their respective microbial populations.
Finally, I plan to address complex questions that will require fine-scale sampling across the chemical and physical gradients that occur within hydrothermal vent systems (i.e., the changes in community structure and diversity associated with these gradients). I plan to continue these analyses with microbial samples retrieved from the Guaymas hydrothermal vents, a high temperature on-axis vent system that is a northern extension of the East Pacific Rise and from Baby Bare, an off-axis vent system near the Juan de Fuca Ridge.

**SIGNIFICANCE OF RESEARCH**

Microbes function at the core of the global ecology. The first priority in microbial ecology in this new era of molecular-based techniques is that of the census-taker and genealogist, so that a foundation can now be set forth for further experimental design and hypothesis testing. The exploration of microbial diversity is important to a sustainable biosphere, because populations as well as communities of microorganisms play critical roles in the degradation and recycling of organic matter, the regeneration and recycling of nutrients and minerals, the production and consumption of gases that affect the global climate, predator-prey relationships, and trophic-level interactions. From a more applied point of view, the exploration of microbial diversity is important because of the roles microbes play in pollutant degradation, the risk assessment associated with the release of genetically engineered microorganisms, the biological control of plant and animal pests, and because they are a virtually untapped natural resource with respect to the potential for discovery of new genes and gene products for use in applications of biotechnology.
REFERENCES


APPENDIX B

PHYLOGENETICALLY ORDERED TAXA LIST (n=106)

Deeply-Rooted Bacterial Phyla

Thermogenic oxygen producers
   *Aquifex pyrophilus*

Thermotogales
    *Thermotoga maritima*

Deinococcus-Thermus subdivision
   *Thermus ruber*
   *Thermus aquaticus*

Planctomycetales
   *Gemmata obscuriglobus*
   *Planctomyces staleyi*

Flexibacter-Cytophaga-Bacteroides phylum
   *Flavobacterium breve*
   *Bacteroides fragilis*
   *Prevotella ruminicola*
   *Flexibacter canadensis*
   *Cytophaga lytica*
   *Cytophaga hutchinsonii*
   *Cytophaga diffluens*
   *Saprospira grandis*
   *Thermonema lapsum*

Spirochete phylum
   *Leptospira interrogans*
   *Serpulina hyodysenteriae*
   *Borrelia burgdorferi*
   *Spirochaeta halophila*
   *Spirochaeta stenostrepta*
   *Treponema pallidum*
Gram-positive Phylum

High G+C subdivision
- *Streptomyces coelicolor*
- *Bifidobacterium bifidum*
- *Arthrobacter globiformis*
- *Propionibacterium innocuum*
- *Corynebacterium otitidis*

Clostridia & Relatives subdivision
- *Clostridium pasteurianum*
- *Clostridium barkeri*
- *Helio bacterium chlorum*
- *Clostridium quercicolum*
- *Selenomonas ruminantium*
- *Desulfitomaculum thermobenzoicum*
- *Desulfotobacterium dehalogenans*
- *Acetogenium kivui*

Bacillus-Lactobacillus-Streptococcus subdivision
- *Streptococcus salivarius*
- *Lactobacillus casei*
- *Bacillus subtilis*

Purple Bacteria Phylum

α-Proteobacteria subdivision
- *Agrobacterium tumefaciens*
- *Brevundimonas diminuta*
- *Rhizobium meliloti*
- *Rhodospirillum rubrum*
- *Azospirillum lipoferum*
- *Azospirillum braziliense*
- *Rhodobacter capsulatus*
- *Rhodopseudomonas blastica*
- *Paracoccus denitrificans*
- *Roseobacter denitrificans*
- *Erythrobacter longus*
- *Nitrobacter winogradskyi*
- *Rhodopseudomonas palustris*
**β-Proteobacteria subdivision**

Comomonas testosteroni  
Alcaligenes xylosoxidans  
Alcaligenes faecalis  
Alcaligenes eutrophus  
Burkholderia cepacia  
Burkholderia pickettii  
Burkholderia solanacearum  
Burkholderia gladioli  
Azoarcus denitrificans  
Rhodocyclylus tenuis  
Spirillum volutans  
Gallionella ferruginea  
Nitrosomonas europaea  
Nitrosococcus mobilis  
Nitrovibrio tenuis  
Leptothrix discophora  
Sphaerotilus natans  
Variovorax paradoxus

**γ-Proteobacteria subdivision**

PVB OTU 4  
Chromatium vinosum  
Thiobacillus hydrothermalis  
PVB OTU 1  
Xanthomonas maltophilia  
Xylella fastidiosa  
Acinetobacter calcoaceticus  
Methylomonas methanica  
Halomonas elongata  
Flavobacterium lutescens  
Pseudomonas flavescentis  
Pseudomonas aeruginosa  
Pseudomonas mendocina  
Vibrio marinus  
Alteromonas ANT-300  
Alteromonas haloplanktis  
PVB OTU 12  
PVB OTU 5  
Vibrio cholerae  
Vibrio fischeri  
Escherichia coli  
Proteus vulgaris  
Aeromonas hydrophila
δ-Proteobacteria subdivision
Desulfurella acetivorans
Desulfovibrio desulfuricans
Syntrophobacter wolinii
Desulfobacter postgatei
Geobacter metallireducens
Desulfoomonile tiedjei
Myxococcus xanthus

e-Proteobacteria subdivision
Helicobacter pylori
Campylobacter jejuni
Thiovulum sp.
PVB OTU 2
PVB OTU 6
Flexispira rappini
Arcobacter nitrofigilis
Wolinella succinogenes
APPENDIX C

SSU rRNA SECONDARY STRUCTURE MODELS

The following pages constitute SSU rRNA secondary structure models for sequences determined and phylogenetically analyzed during this dissertation.
Secondary Structure: small subunit ribosomal RNA

OTU 2 - Pele's Vents

Domain: Bacteria
Kingdom: Purple Bacteria
Order: epsilon
Secondary Structure: small subunit ribosomal RNA

OTU 1 - Pele's Vents

Domain: Bacteria
Kingdom: Purple Bacteria
Order: gamma
Secondary Structure: small subunit ribosomal RNA

OTU 4 - Pele's Vents

Domain: Bacteria
Kingdom: Purple Bacteria
Order: gamma
Secondary Structure: small subunit ribosomal RNA

OTU 5 - Pele's Vents

Domain:  Bacteria
Kingdom:  Purple Bacteria
Order:    gamma

145
Secondary Structure: small subunit ribosomal RNA

OTU 12 - Pele's Vents

Domain: Bacteria
Kingdom: Purple Bacteria
Order: gamma
Secondary Structure: small subunit ribosomal RNA

OTU 10B - Pele's Vents

Domain: Bacteria
Kingdom: Purple Bacteria
Order: delta
Secondary Structure: small subunit ribosomal RNA

OTU 9A - Pele's Vents
Domain: Bacteria
Affiliation: Paraphyletic assemblage
Secondary Structure: small subunit ribosomal RNA

ANT-300

Domain: Bacteria
Kingdom: Purple Bacteria
Order: gamma
Secondary Structure: small subunit ribosomal RNA

OTU 2 - Pele's Vents

Domain: Archaea
Kingdom: Crenarchaeota
Affiliation: Marine Group I
Secondary Structure: small subunit ribosomal RNA

OTU 1 - Pele's Vents

Domain: Archaea
Kingdom: Euryarchaeota
Affiliation: Marine Group II