DISTRIBUTIONS AND DIVERSITY OF AEROBIC ANOXGENIC
PHOTOTROPHIC BACTERIA ALONG ONSHORE OFFSHORE TRANSECTS NEAR
PACIFIC ISLANDS

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF
HAWAI'I IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF
Masters of Science

IN
OCEANOGRAPHY

AUGUST 2008

By
Anna E. Ritchie

Thesis Committee:
Zackary Johnson, Chairperson
Susan Brown
Matthew Church
We certify that we have read this thesis and that, in our opinion, it is satisfactory in scope and quality as a thesis for the degree of Master of Science in Oceanography.

THESIS COMMITTEE

[Signatures]

Chairperson

[Name]

[Name]
This thesis is dedicated to Jan, Paul, Drew, Heidi, and Baby Ritchie.
Acknowledgments

I would like to thank the crew of the R/V Kilo Moana for assistance during the WP2 cruise and mini-WP2 cruise. Also Kathy Barbeau, crew of the R/V New Horizon, and Yajuan Lin for assistance during the California DCM1 cruise. I am grateful to Dan Reineman, Dan Barschis, and Eric Bachman for assistance during sampling in Kaneohe Bay. Also matlab codes for Kaneohe Bay and graphics help from Jerome Aucan, Thomas Decloedt, and Michael Romano. Thanks to NSF for financial assistance.
Abstract

Aerobic anoxygenic phototrophic bacteria (AAP's) population abundances are not well characterized, in spite of their potential importance for the microbial loop and primary production. AAPs have a photoheterotrophic metabolism, thus both inorganic and organic nutrients, as well as physical variables likely drive their ecological distributions. To explore AAP abundances and the environmental variables that may be regulating them, we quantified AAP abundance by counting the pufM biomarker along onshore/offshore transects near 5 distinct islands in the Pacific Ocean (Oahu, Molokai, Futuna, Aniwa, and Lord Howe) with steep environmental gradients. Abundance patterns are further explored by investigating the genetic diversity of the AAP community using pufM clone libraries and QPCR dissociation curves. Overall we report small but sometimes significant AAP populations that increase near shore and are comprised of distinct genetic clades.
Table of Contents

Abstract ........................................................................................................ v
List of Tables ................................................................................................ viii
List of Figures ........................................................................................... ix
Chapter 1: Introduction .............................................................................. 1
Chapter 2: Aerobic Anoxygenic Phototrophic Bacterial Abundance ........... 4
  Methods ........................................................................................................ 7
  Sampling ...................................................................................................... 7
  DNA Extraction ....................................................................................... 8
  QPCR Standards ...................................................................................... 9
  QPCR Enumeration of AAPs .................................................................. 10
  QPCR Negative Controls ...................................................................... 11
  Enumeration of Total Bacteria and Picophytoplankton using Flow Cytometry .................................................. 12
  Nutrient Analyses ................................................................................... 12
  Statistical Analyses ................................................................................. 13
Results .......................................................................................................... 13
  AAP Abundance ..................................................................................... 13
  Environmental Variables: Salinity and Temperature ...................................... 15
  Environmental Variables: Chlorophyll, Photosynthetic Picoplankton, and Total Bacteria ................................. 15
  Environmental Variables: Nutrients .......................................................... 16
  Environmental Variables Correlated with AAPs Abundance ...................... 16
Discussion ...................................................................................................... 17
Chapter 2 Figures and Tables ..................................................................... 22
Chapter 3: Aerobic Anoxygenic Phototrophic Bacterial Diversity ............ 30
  Methods .................................................................................................... 32
  Sampling .................................................................................................. 32
  Preliminary AAP’s Diversity Study at Kaneohe Bay .................................... 33
  PufM Clone Libraries ............................................................................... 34
  Comparing Clone Libraries using SONS and Confirming Clone Library Saturation ......................................................... 35
  Examining Diversity of AAPs using Dissociation Curves ........................... 35
Results .......................................................................................................... 36
  Kaneohe Bay Preliminary Diversity Study Results ..................................... 36
  Clone Libraries Results ........................................................................... 37
  Dissociation Curves of Selected Clone Library Sequences .......................... 39
  Clone Library Community Structure ...................................................... 41
  Near shore and Offshore Diversity and Community Structure .................... 41
  Broad AAP Diversity Determined by Dissociation Curves ......................... 43
Discussion .................................................................................................................. 44
Chapter 3 Figures and Tables .................................................................................. 48
References ................................................................................................................... 56
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. List of Locations and Chlorophyll Size Fractions</td>
<td>28</td>
</tr>
<tr>
<td>2. List of Correlation and Model II Linear Regression Results</td>
<td>29</td>
</tr>
<tr>
<td>3. List of SONS Results for Aniwa, California, Oahu, and Onshore/Offshore Libraries</td>
<td>55</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.</td>
<td>Map of Sampled Locations</td>
<td>22</td>
</tr>
<tr>
<td>2.</td>
<td>QPCR AAP Standard Curve</td>
<td>23</td>
</tr>
<tr>
<td>3.</td>
<td>Depth Profiles for AAP Concentrations</td>
<td>24</td>
</tr>
<tr>
<td>4.</td>
<td>AAP Concentrations Versus Distance from Shore for Pacific Islands and California Transects</td>
<td>25</td>
</tr>
<tr>
<td>5.</td>
<td>AAP Concentrations Versus Distance from Shore for Kaneohe Bay Transects</td>
<td>26</td>
</tr>
<tr>
<td>6.</td>
<td>Positive Correlations of AAP Concentrations with Chlorophyll, <em>Synechococcus</em>, and Bacteria</td>
<td>27</td>
</tr>
<tr>
<td>7.</td>
<td>Clone Libraries Dissociation Curves</td>
<td>48</td>
</tr>
<tr>
<td>8.</td>
<td>Neighbor-joining Trees for Aniwa, California, and Oahu Libraries</td>
<td>49</td>
</tr>
<tr>
<td>9.</td>
<td>Rarefraction Curves</td>
<td>50</td>
</tr>
<tr>
<td>10.</td>
<td>Neighbor-joining Tree for Onshore and Offshore Sequences</td>
<td>51</td>
</tr>
<tr>
<td>11.</td>
<td>Neighbor-joining Tree for Onshore and Offshore Sequences With Chlorophyll, <em>Synechococcus</em>, and Bacteria</td>
<td>52</td>
</tr>
<tr>
<td>12.</td>
<td>Diversity of AAP Community for Aniwa, Futuna, and Kaneohe Bay Transects</td>
<td>53</td>
</tr>
<tr>
<td>13.</td>
<td>Diversity of AAP Community for California, Lord Howe, Oahu, and Molokai</td>
<td>54</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Microbial oceanography is a field that focuses on understanding the distribution, community composition, and interactions of bacteria, viruses, archaea, and protists with each other and the environment. The terminology “microbial” is an interchangeable term in microbial oceanography literature; here the term “microbial” will refer to the marine bacterial community.

Microbial oceanography dates back to 1677 with Antonie van Leeuwenhoek’s description of small animals in a drop of water (Van Leeuwenhoek, 1677 referenced in Karl, 2007). In recent years this field has exponentially expanded due to technological advancements such as genetics and flow cytometry to enumerate and identify key players in the microbial community that elude cultivation. In the past thirty years scientists have identified marine microbes as an alternative pathway to facilitate recycling of nutrients by utilizing DOM, then being grazed upon by protists; this is known as the microbial loop (Pomeroy, 1974, Azam et al., 1983). The importance of marine microbes can vary from very small scales of individual microbial interactions to the global scale. On the global scale marine microbes influence the degradation and production of DMS, which is an important greenhouse gas more potent than carbon dioxide in its control on global climate (Andreae et al., 1997, Zemmelink et al., 2005). While some microbes directly impact the global climate, some such as Prochlorococcus can be used as a model to help us understand what primary production could be like if the oceans warmed, stratification increased, and waters become nutrient depleted due to anthropogenic induced global warming (Behrenfeld, et al., 2006). Almost daily our perceived knowledge of how the
ocean and the microbial community functions is challenged and expanded to new dimensions.

While these examples are a few of the periodic large discoveries of the field there are many other achievements addressing the current questions and problems of microbial oceanography. The driving questions of our field are: How abundant are marine microbes? What is the diversity of marine microbes? What are the community’s metabolic capabilities? How do microbes interact and how are they distributed through the world’s ocean? An additional goal in this field is to link microbial communities to geochemical cycles and physical processes in the ocean.

The first order questions of how abundant and who makes up the microbial community are one of the most challenging questions due to the sheer abundance of marine bacteria. One approach is to investigate a part of the community using biomarkers unique to a given population. A few examples of commonly used biomarkers are pigments, functional genes, or 16S rDNA. Genetic biomarkers describing the diversity of marine microbes has advanced dramatically but is limited by our ability to only search for what we already know genetically exists. Therefore much of microbial diversity is uncharacterized and leaves scientist questioning the feasibility of determining the diversity of marine bacteria (O’Malley, 2007, Whitfield, 2005, Pedros-Alio, 2006, Pommier et al., 2007, Longhurst, 1998). While knowing how many different types of bacteria are present in the ocean would help us to comprehend how feasible it is to describe microbial diversity, it is not the only important aspect of understanding microbial diversity. A majority of what we have learned about marine microbial diversity
has been through non-cultivated techniques due to the inability to culture up to 99% of marine bacteria in the ocean. While molecular information has widened our understanding of the types of microorganism in the sea, we must keep in mind this molecular information must be put into ecological context through cultivation of the organisms (Giovannoni and Sting, 2005). For example, cultivated representatives and environmental studies of *Prochlorococcus* strains MED4 and MIT9312 have revealed different horizontal and vertical distributions within the ocean while their genetic similarity for the 16S rRNA would suggest they are the same strain (Partensky et al., 1999). The linking of diversity data and bacterial distributions with cultured strains proves to be critical information for the forward movement of this field.

The following chapters use a genetic biomarker to describe the diversity and distribution of a group of marine bacteria termed aerobic anoxygenic phototrophic bacteria (AAP’s). AAPs are a group of mixotrophic bacteria containing the pigment bacteriochlorophyll a (Bchl a). In chapter 2, I describe the distributions of AAPs along onshore and offshore transects near distinct Pacific Islands using quantitative PCR targeting the functional gene *pufM*. The *pufM* gene encodes for a reaction centre associated with the bacteriochlorophyll a photosynthetic system. In chapter 3, I use this same functional gene to investigate the diversity of AAPs along the same onshore and offshore transects. The overall purpose of these two chapters is to understand AAPs ecological distribution through their abundance, diversity, and the environmental variables which drive these patterns.
Chapter 2: Aerobic Anoxygenic Phototrophic Bacterial Abundance

AAPs are a group of mixotrophic bacteria which contain the pigment bacteriochlorophyll a (Bchl a). Oxygen, light, and organic carbon sources are known to influence the formation of this pigment (Yurkov and Beatty, 1998a). Generally AAPs have low Bchl a concentrations per cell (Kolber et al., 2001) and previous research suggests they cannot survive on a sole photoautotrophic metabolism (Yurkov and Beatty, 1998a). AAPs have multiple metabolic pathways that may confer an advantage in inhabiting a variety of environments or environments in constant flux. Laboratory experiments and whole genome sequencing have revealed a wide range of metabolic capabilities such as utilization of low weight organic carbon, growth optimization on complex organic carbon, nitrification, carbon dioxide fixation, light enhanced cellular growth, preferred suboxic environments, aggregate formations, and high amounts of carotenoids for UV protection (Fuchs et al., 2007, Denner et al., 2002, Gich and Overmann, 2006, Qiang et al., 2006). Due to the broad metabolic capabilities of AAPs it is expected and has been found that they live in a variety of environments ranging from freshwater, marine environments, soil, hot springs, Antarctic lakes, and hydrothermal vents (Yurkov and Beatty, 1998a, Rathgeber et al., 2004, Labrenz et al., 2005).

AAPs anoxygenic phototrophy is thought to have been obtained from their anaerobic counterparts (Kolber, 2007). This form of phototrophy uses reduced sulfur, iron, or unknown electron donors and does not produce oxygen as a byproduct. The lack of oxygen evolution is one of the main reasons AAPs have not been included in some estimates of photosynthesis. In addition, AAPs have not been included in estimates of
photosynthetic biomass because they do not contain chlorophyll. Therefore, AAPs have historically been included within heterotrophic bacteria community and bacterial production measurements. Photoheterotrophy within the bacterial community has been shown to be important in that light stimulated leucine uptake increased significantly compared to conventional methods of measuring bacterial production in the dark (Church et al., 2004). Follow up studies on this phenomenon reported that up to 40% of the increase could be due to unidentified photoheterotrophs such as AAPs (Michealou et al., 2007). These results suggest that photoheterotrophy is a common strategy that some current measurements of photosynthesis and respiration do not resolve.

In addition to understanding what proportion of phototrophy is carried out by AAPs, another major goal is to understand how this new form of phototrophy affects the carbon cycle. While the abundance of AAPs within the oligotrophic ocean suggests they do not contribute significantly to the carbon cycle, other AAP characteristics such as growth rates suggest AAPs contribution can be significant. For example, Koblizek and collaborators (2007) predicted AAPs would account for 4-50% of total bacterial production in the oligotrophic, if growth rates were 1 division per day for AAPs and an abundance of 2-4% of the total bacterial community. Thus I would argue other AAP characteristics such as growth rate, amount of Chl a per strain, and cell size should be considered before concluding they are not important contributors to the carbon cycle.

Understanding the ecology of AAPs will facilitate answering these bigger issues discussed above. The current study is focused on describing the distribution of AAPs along onshore/offshore transects and determining the relationship with environmental
variables driving their ecology. Our current understanding of the distribution of AAPs is that they are most abundant near coastal waters, ranging from 0.5% to 34% of the total community (Zhang and Jiao, 2007, Du et al., 2006, Cottrell et al., 2006, Sieracki et al., 2006, Waidner and Kirchman, 2007, Schwalbach and Fuhrman, 2005). In the oligotrophic ocean abundance estimates are usually less than 1 to 6% of the total community (Cottrell et al., 2006, Jiao et al., 2007, Yutin et al., 2007) excluding two studies that estimate AAPs to be as much as 11% (Kolber et al., 2000) and 24% (Lami et al., 2007) of the total community. These periodic reports of high AAP abundance in the oligotrophic ocean challenge our understanding of what controls the distribution of AAPs, and prompted this study near Pacific Islands. Here I use gradients in a variety of physical and chemical variables of these Pacific Islands to determine what are the proximal variables responsible for determining AAP abundance. I hypothesized that controls on AAPs distributions along natural environmental gradients could be light, nutrients, salinity, and photosynthetic community composition. I report that AAPs generally comprise 1% of the total bacterial community, except for the most eutrophic location where they accounted for more than 4% of the total bacterial community. The abundance of AAPs was most closely correlated with chlorophyll (photosynthetic biomass), *Synechococcus*, and total bacteria.
METHODS:

**Sampling**

Eight onshore/offshore transects were conducted in the Pacific Ocean at the following locations: California (28° 36'N, 125° 4'W), Aniwa (18° 10'S, 169° 15'E), Lord Howe (31° 33'S, 159° 05'E), Futuna (19° 32'S, 170° 13'E), Oahu (21° 28'N, 157° 59'W), Molokai (21° 08'N, 157° 02'W), and two in Kaneohe Bay located on the island of Oahu (21° 27' 35"N, 157° 48' 15" W). Surface water samples were collected in dark amber bottles for chlorophyll, salinity, flow cytometry, nutrients (NO₂, NO₃, PO₄, and NH₄), and DNA for abundance measurements. Two of the onshore/offshore transects located in Kaneohe Bay were sampled three times (June 13th 2006, August 4th, 2006, and June 21st 2007) (Figure 1). In addition to surface samples, two depth profiles were measured to explore depth distribution of AAPs. Depth profiles were conducted off of the island of Oahu and sampled using oceanographic niskin bottles mounted to a CTD equipped rosette.

Along Aniwa, Futuna, and Kaneohe Bay transects, temperature was recorded using a digital thermometer (Traceable). Temperature and salinity measurements for California, Oahu, and Molokai were made with a CTD mounted on the rosette. Kaneohe Bay salinity measurements were made in duplicate using a refractometer (Vista).

Total (0.22 μm) or size fractionated (0.8, 2.0, or 5.0 μm filters, see Table 1 for details) chlorophyll measurements were made by filtering 100 ml of sample onto polycarbonate filters (GE Water and Process Technologies) under low filtration pressure (< 15 mm Hg). Filters were extracted in 100% methanol for 24 hours in the dark at -4 °C,
and then read on a Turner 10AU fluorometer calibrated according to Porra et al. (1989). Chlorophyll a measurements were calculated following Welschmeyer (1994).

DNA was collected by filtering 50 or 100 ml of water through 0.22 μm polycarbonate filter (GE Water and Process Technologies) followed by 3 ml of preservation solution (10mM Tris,100mM EDTA,0.5 NaCl at a pH of 8.0). Filters were placed in bead beating tubes and stored at -80 °C freezer until extraction.

**DNA Extraction**

DNA was extracted from duplicate filters using physical and chemical procedures following the Gentra DNA extraction kit, with the following modifications: Briefly, bead beating tubes containing filters were removed from -80 °C freezer and 0.25 g of sterile 0.1 mm zirconium beads (Biospec) were added. Seven hundred fifty microliters of cell lysis solution (Gentra systems) was added to each tube and samples were put on ice for 1 minute, then placed in bead beaters for 1 minute at 4800 RPM's. Tubes were incubated for 5 minutes at 80 °C on a heating block, after which 4 μl of RNAse A was added and tubes were inverted 25 times before incubating for 30 minutes at 37 °C. After incubation tubes were cooled on ice for 1 minute before adding 250 μl of Protein Precipitation solution (Gentra Systems) and vortexed for 20 seconds at a maximum speed of 3200 RPM's (VWR). Tubes were placed on ice for 5 minutes before centrifuging at 16,000 x g for 3 minutes. Nine hundred microliters of the supernatant were transferred to a new 1.5 ml microcentrifuge tube and centrifuge as above. Seven hundred fifty microliters of the supernatant was transferred to a new microcentrifuge tube containing 750 μl of 100%
isopropanol. Samples were inverted 50 times before centrifuging for 1 minute at 16,000 x g. Supernatant was poured off and 300 µl of 70% ethanol was added and tubes were inverted several times. Centrifuging was repeated for 1 minute at 16,000 x g. Afterwards supernatant was poured off and tubes were allowed to dry in a laminar flow hood for 15 minutes before 100 µl of DNA Hydration Solution (Gentra Systems) was added. Tubes were stored overnight at room temperature to allow DNA to rehydrate and stored in -80 °C freezer.

**QPCR Standards**

Four sets of QPCR standards were made from quadruplicate cultures of *Erythrobacter longus* Strain NJ3Y (Koblizek et al. 2003) cultivated in rich media (F/2 media (Guillard and Ryther, 1962, Guillard, 1975), 0.5 g peptone/L, and 0.1 g yeast extract/L) at 30 µmol quanta/m²/sec on a 12/12 light cycle. Cultures were grown until the late log phase of growth and a maximum concentration of at least 1 x 10⁹ cells/ml as determined by epi fluorescent microscopy (Nikon) of DAPI (Sigma Aldrich) stained cells (Porter and Feig Y., 1980), and flow cytometric (FASCAliber) cell counts of SYBR green (Invitrogen) stained cells. Cultures were diluted with sterile seawater collected from Station ALOHA (22° 45' N, 158° 00'W) and DNA was extracted as described above. A serial dilution of the extracted DNA was performed to create standards that ranged between 1 x 10⁷ to 1 x 10⁻¹ cells/ml.

The average for all QPCR threshold cycle values (Cₚ) at a given *pufM* standard concentration were calculated and plotted against the known *pufM* standard concentrations (Figure 2). A model I linear regression resulted in a slope of -2.13, which
supports near a 100% amplification efficiency. The average r-squared value for standard curves from all QPCR runs was 0.99. The standard detection limit was $3.7 \times 10^1$ \textit{pufM} copies/ml, which was determined by the lack of resolution between the lowest standard and the no template control.

\textit{QPCR Enumeration of AAPs}

Enumeration of AAPs was achieved using primers specific to the functional gene \textit{pufM}, which encodes a central photosystem reaction center protein that is unique to AAPs. The abundance of AAPs was calculated assuming one copy of \textit{pufM} per genome. This is consistent with all genomic data available, including the AAP strain \textit{Roseobacter dinitrificans} (Swingley et al., 2007). Primers (\textit{pufM557F} (5'TAC GGS AAC CTG TWC TAC 3') and \textit{pufMW}AWR (5'GCR AAC CAC CAN GCC CA 3')) amplified 240 base pairs of the \textit{pufM} gene (total \textit{pufM} gene is approximately 1000 bp) as described by Yutin et al. (2005) and Waidner and Kirchman (2007).

Duplicate QPCR reactions were analyzed for each standard dilution, environmental DNA extract, and for all negative controls. For each station a total of four QPCR reactions were performed (duplicate QPCR reaction per filter plus duplicate filters per station). Each 25 μl QPCR reaction contained the following: 2.5 μl of DNA template, 1X SYBR iTaq supermix (BioRad), and 0.1 μmol L\textsuperscript{-1} of forward primer \textit{pufM557F} and reverse primer \textit{pufm_WAWR}.

QPCR assay was performed following Kirchman and Waidner (2007) PCR protocol (95 °C for 10:00 min, then 40 cycles of 95 °C for 15 sec, 56 °C for 45 sec, 72 °C
for 45 sec) using a 7300 Real time PCR system (Applied Biosystem). Enumerations were calculated using manufacturer’s software to create a linear regression of fluorescence for each standard versus the known concentration of cells determined by flow cytometry.

**QPCR Negative Controls**

In my original evaluation of various suppliers SYBR-Green QPCR mastermixes, I observed significant non-specific amplification of non-target template DNA using \( pufM \) primers. As a result, I used the following negative controls to monitor non-specific binding during the QPCR assay: water, *Altermonas* UH00601, and *Prochlorococcus* MIT9312. Each negative control was extracted similar to environmental samples and standards. Negative controls were extracted at \( 10^6 \) cells/ml concentration as determined by DAPI stained epi-fluorescent microscopy.

Negative controls concentrations (UH06001 or MIT9312), which gave statistically similar results, were averaged together for all QPCR assays. Negative controls that were added at \( 1 \times 10^6 \) cells/ml, resulted in amplification equivalent to 276 \( pufM \) copies/ml for UH06001 and 225 \( pufM \) copies/ml for MIT9312. Therefore, amplification of negative controls was four orders of magnitude below the concentrations added to each reaction.
**Enumeration of Total Bacteria and Picophytoplankton using Flow Cytometry**

Flow cytometry samples were preserved by collecting 1 ml of sample water in a cryovial containing 6 µl of 25% glutaraldehyde (Tousimis) for a final concentration of 0.125% . Samples were mixed and placed in the dark for 10 minutes before storing in -80 °C freezer or liquid nitrogen.

A FASCAlber flow cytometer equipped with a syringe pump (flow rate 10 µl/min) collected the following variables: forward light scatter, side light scatter, 530/30 nm, 670 nm LP, and 585/42 nm. *Prochlorococcus*, *Synechococcus*, and small pico eukaryote populations were enumerated based on unique combinations of the 5 variables following Olson et al. (1990). Total bacteria were processed similarly, but stained with 1X Sybr Green I (Molecular Probes) for 30 minutes in the dark before processing (Marie et al., 1997). All samples were run with 1 µm yellow-green microspheres (Polysciences, Inc.) as an internal standard.

**Nutrient Analyses**

Nutrient samples for Kaneohe Bay, Oahu, Molokai, Aniwa, Lord Howe, and Futuna were processed following Hernandez-Lopez and Vargas-Albores (2003). Nitrite standard curves were made up of 17 concentrations ranging between 0.7 - 50 µM. Nitrate standard curves were made up of 17 concentrations ranging between 0.7 – 40 µM. Phosphate standard curves were made up of 17 concentrations ranging between 0.5 – 30 µM. Samples were analyzed in triplicate for all measurements except for phosphate for the islands of Aniwa, Oahu, Lord Howe, and Futuna, which were analyzed in duplicate.
Nutrient concentrations were calculated using model I linear regressions fitted to the standard concentrations against absorbance values. Based on environmental nutrient concentrations the ranges of standards were restricted to 15 µM as the highest standard. This provided a better linear regression with r-squared values ranging between 0.98-0.99 for nitrate, nitrite, and phosphate. The limit of detection for nitrite, nitrate, and phosphate nutrient assays were 1 µM.

Statistical Analyses

Statistical calculations such as correlations, model I regressions, and non-parametric calculations were made using MINITAB 14. Model II regressions were calculated using a Matlab code created by Edward T. Peltzer. Plots of the island transects were accomplished with Matlab. Images of Futuna, Aniwa, Lord Howe, Oahu, California, and Molokai were obtained from Google Earth then digitized with the z graph digitizer 1.9. Tables of latitude and longitude obtained from the digitized image were then imported into Matlab to graph the islands. Kaneohe Bay plots were formed in Matlab via high resolution bathymetry data supplied by Dr. Jermone Aucan.

RESULTS:

AAP Abundance

Two vertical profiles of AAPs concentrations were measured at stations along the Oahu transect (Figure 3). AAP abundance was highest at the surface and decreased with depth (Figure 3). AAPs were detectable to about 100 meters.
Most island and California surface samples had AAP concentrations that were less than $6 \times 10^3$ cells/ml and less than 1% of the total bacterial community (Figure 4). Kaneohe Bay had the highest concentration of AAPs at $1.35 \times 10^5$ cells/ml (Figure 5A). Kaneohe Bay also had the highest AAP's percentage (4.3%) of the total bacterial community (Figure 5B). Offshore stations in Kaneohe Bay had AAP concentrations similar to those reported for samples collected offshore in California and the other island transects.

Most locations had the highest AAP concentrations near shore. AAP abundances at the offshore stations were consistently lower than those observed in the near shore stations. Although there was significant variability between locations, the overall patterns of AAP abundance and as a percentage of the total bacterial community show the highest values near shore.

Variability seen in the overall offshore AAP abundance pattern can be explained by analyzing each location/transect separately. For Aniwa, California, and all of Kaneohe Bay transects a decrease in AAP abundance towards the offshore stations was observed. AAP abundance along the Molokai and Futuna transects generally increased with distance from shore. AAP abundance along the Oahu and Lord Howe transects varied little between stations (Figure 4).

While no seasonal study was performed in this study it is surprising the consistency of AAP abundance between the three sampling dates in Kaneohe Bay. Enumeration of AAPs for Kaneohe Bay's north and south channel transects were similar
between sampling dates. While the total bacterial estimates were less consistent, resulting in change in the AAP to total bacterial abundance (Figure 5).

**Environmental Variables: Salinity and Temperature**

Salinity ranged from 31 to 38. Temperature ranged from 17.4 °C (California) to 28 °C (Futuna and Aniwa). Temperatures varied little along the onshore/offshore transects with the exception of Kaneohe Bay (6.21.07) where temperatures ranged between 21.9 °C to 26.6 °C. For a majority of the transects high salinity and lower temperatures were seen offshore, while lower salinity and higher temperatures were seen near shore.

**Environmental Variables: Chlorophyll, Photosynthetic Picoplankton, and Total Bacteria**

Total chlorophyll a was greatest in Kaneohe Bay (6.13.06) at 3.1 µg/L and (8.4.06) 2.8 µg/L. Of the other locations, the highest chlorophyll a concentration was 0.5 µg/L (California transect). Chlorophyll a tended to decrease as distance from shore increased, similar to patterns observed for AAP abundance. Total bacteria concentrations ranged from 3.2 x 10^6 to 1.26 x 10^5 cells/ml with an average of 9.4 x 10^5 (+/- 5.3 x 10^5 cells/ml). *Prochlorococcus* abundances varied between 2.4x 10^3 to 1.2 x 10^5 cells/ml. *Prochlorococcus* composed less than 1 to 28 percent of the total bacterial community. *Synechococcus* abundances varied between 7 x 10^1 to 6.7 x 10^5, comprising less than 1 to 21.9 percent of the total bacterial community. *Synechococcus* concentrations were greatest in Kaneohe Bay and lowest for Futuna, Aniwa, Oahu, and Molokai.
Picoeukaryotes were quantified by flow cytometry analysis, but our flow cytometry setting cannot capture the whole community while simultaneously enumerating *Prochlorococcus*, as a result picoeukaryotes were excluded from this analysis.

Correlations and Model II linear regressions were calculated to assess whether distance from shore, a proximity for natural environmental gradients, explained the variability in *Prochlorococcus, Synechococcus*, and total bacterial concentrations. Distance from shore did not explain any variability for *Prochlorococcus* ($r = 0.193, p = 0.293, n=77$), *Synechococcus* ($r = -0.209, p = 0.065, n =77$), and total bacteria ($r =-0.12, p =0.293, n =77$) concentrations.

**Environmental Variables: Nutrients**

Phosphate and nitrite concentration were below detection limits (1 μM) for all samples. Nitrate was detectable in 25 out of 59 samples at concentrations ranging from 1μM to 3.3 μM. No significant relationships were found with nitrate and *Prochlorococcus, Synechococcus*, or total bacteria concentrations.

**Environmental Variables Correlations with AAPs Abundance**

Model II linear regressions were calculated to assess the relationship between environmental variables and AAP abundances (Table 2). In spite of significant variability in environmental variables, no relationships were found between AAP cells/ml and distance from shore (nm) ($r =-0.181 , p =0.111, n =79$), nitrate ($r =0.03, p =0.965, n =25$), chlorophyll greater than 2 μm ($r =0.0518, n =77$) and *Prochlorococcus* ($r =-0.220, p$
= 0.033, n = 94). However, AAPs were correlated with total bacteria ($r = 0.784$, $p \leq 0.005$, $n = 94$), *Synechococcus* ($r = 0.871$, $p \leq 0.005$, $n = 94$), total chlorophyll ($r = 0.802$, $p \leq 0.005$, $n = 94$), and 0-2 μm chlorophyll a ($r = 0.743$, $p \leq 0.005$, $n = 77$). In general, the relationships between AAP abundance and *Synechococcus*, total bacteria, 0-2 μm chlorophyll a, and total chlorophyll were largely driven by the variability observed in Kaneohe Bay. At the other sampling sites, the regressions were consistent with the patterns observed in Kaneohe Bay, but the low variability in AAPs and environmental variables reduced the statistical power of comparison (Figure 6).

Outliers (2 standard deviations above the mean) were identified by box and whisker plots and removed from the analyses to determine if the linear regression were heavily influenced by extreme values. Further statistical analyses were performed on positive relationships to confirm that linear regression assumptions of normality were met. Transformations and non-parametric analysis were also made. Data was also fourth root transformed and model II regressions were recalcualted. For all of these analyses, the $r$ values changed slightly but the relationships were still consistent. The Spearman’s rank correlation non-parametric analysis was performed.

**DISCUSSION:**

Overall AAP concentrations were highest near shore, particularly in Kaneohe Bay where concentrations were greatest near the river input to the central part of the bay. Kaneohe Bay has two river inputs (southern and central bay), neither of which significantly affected the salinity at the sampling site, but the central bay was visually
more turbid and murky. AAPs have been reported to be influenced by river inputs due to increased particle load, decreased light attenuation, freshwater source of AAPs, and changes in salinity (Waidner and Kirchman, 2007). For example, Waidner and Kirchman’s study (2007) of the Delaware and Chesapeake estuaries reported the highest AAP concentrations at the turbidity maximum and AAP concentrations were correlated to light attenuation and particle associated AAPs. While my study did not measure light attenuation or particle associated AAPs, I can expect that our high AAP concentrations near these river inputs were influenced by light and particle associated AAPs.

Kaneohe Bay was the only location with a documented river source, so I also considered other environmental characteristics such as benthic habitat and coastal water residence time as factors controlling AAP bacterial variability. Interestingly, a common trend appeared when grouping islands into those with lagoons protected by barriers reefs (Kaneohe Bay, Lord Howe, and Aniwa) and those with steep cliffs and fringing reefs (Futuna and Molokai) (Whittaker, 1998). AAP concentrations decreased with distance from shore for islands with barrier reefs, in contrast islands with steep cliffs had no patterns or had the highest concentrations furthest from shore. These trends maybe linked to benthic habitats that influence the residence time of coastal waters and circulation patterns.

While there were some weak relationships between AAPs and the physical variables, the strongest correlations were observed for biological variables. Positive relationships were documented for total chlorophyll, *Synechococcus*, total bacteria, and the 0 – 2 µm chlorophyll fraction (Table 2 and Figure 6). Positive relationships between
total chlorophyll and AAP abundance have been well documented since Kolber and collaborators (2000) open ocean description of AAPs. The actual mechanisms as to why AAPs are correlated with total chlorophyll are unknown. One current idea is that AAPs are similar to heterotrophic bacteria, whose relationship to chlorophyll is tightly coupled due to the dependence on DOM derived from leaky phytoplankton cells. Another possibility is AAPs have similar requirements as phytoplankton, therefore are in direct competition for these resources; hence the co-occurrence.

DOM derived from phytoplankton is complex and biological communities that utilize freshly produced DOM vary depending on DOM pool constituents (Teira et al., 2008, Hold et al., 2001, Schafer et al., 2002). Experiments by Pinhassi and collaborators (2004) demonstrated that dominance by a specific species of phytoplankton can alter bacterioplankton community composition due to the availability of preferred DOM sources. The most eutrophic site Kaneohe Bay was dominated by *Synechococcus* and small picoeukaryotes (Cox et al., 2006), whereas all other locations with concentrations of AAPs at 1% or less were dominated by *Prochlorococcus*. While it is difficult to infer if AAPs are influenced by DOM produced from different picocyanobacterial communities, we do know that AAPs will be influenced by the composition of the DOC pool due to their dependence on organic carbon for growth.

Of all the relationships reported here I find that the most revealing and exciting was the positive relationship between AAPs and *Synechococcus*. *Synechococcus* is one of the most abundant picocyanobacteria in the ocean and peak abundances are observed near coastal upwelling regions, as well as periodically in oligotrophic waters.
(Zwirglmaier et al., 2008, Zwirglmaier et al., 2007). Similar patterns have been described for AAPs, and suggest that similar mechanisms are controlling these two bacterial groups. In addition to similar distribution patterns, whole genome analysis and QPCR/ fish probe targeted *Synechococcus* ecotypes reveals that genetically different types of *Synechococcus* are capable of using phototrophic and heterotrophic metabolisms (Palenik et al., 2006). The mechanisms that control the whole *Synechococcus* genus as well as the different ecotype distributions are still unknown. I speculate that better understanding of *Synechococcus* distributions could reveal what controls AAP distributions due to their shared resources and genetic evidence of open ocean and coastal types.

In conclusion, I have reported Pacific Island transects with AAP estimates similar to oligotrophic locations and other volcanic islands (Schwalbach and Fuhrman, 2005). The low nutrients, chlorophyll a concentrations, and the dominance of *Prochlorococcus* at these sites suggest they are oligotrophic environments, which is not surprising due to their remote locations within the Pacific Ocean. Low abundance of AAPs reported here, and in other studies, suggests AAPs play a less critical role in the oligotrophic ocean carbon cycle. However, even at < 1% of the total bacterial population, AAPs may still contribute significantly. Further research is needed on AAP bacterial growth rates, strain specific amount of Bchl a pigment, variables affecting photosynthetic efficiency, and susceptibility to grazing. In contrast to these oligotrophic locations, I was able to sample Kaneohe Bay that represented a mesotrophic body of water, and AAP estimates increased with increasing total chlorophyll a. While this relationship is not new (Kolber et al.,
2001) it raises the question of AAPs importance in coastal waters. Previous studies report that AAP's comprise up to 34% of the total bacterial community, but their importance are not considered important to coastal water carbon cycling due to the dominance of large phytoplankton (Moran, 2007).
FIGURE 1: Locations sampled within the Pacific Ocean. A.) Island of Oahu and Molokai with an expanded view of Kaneohe Bay. B.) California transect C.) Island of Futuna D.) Island of Aniwa E.) Lord Howe Island
Figure 2: Average cycle threshold ($C_t$) for all QPCR assays versus AAP standard concentrations determined by flow cytometry. Average standard detection limit (---). Standard deviation bars are +/- one deviation.

$y = -2.2(\ln(x)) + 48.2$ with $R^2 = 0.99$. 
Figure 3: Depth profiles of AAP abundance determined by QPCR of the *pufM* gene at Oahu Stations OA1 (diamonds) and OA4 (squares). A.) AAP cells/ml. Dashed black line represents the no template control. Standard deviations are +/- one deviation. B.) % AAP/Bacteria.
Figure 4: AAP abundance for Aniwa (○), Futuna (●), Lord Howe (▲), Molokai (▲), Oahu (★), and California (★) estimated by QPCR of the pufM gene. 
A.) AAP cells/ml. Orange line indicates no template control. Standard deviations are +/- one deviation. B.) % AAP/Bacteria
Figure 5: AAP abundance estimated with QPCR of the *pufM* gene for Kaneohe Bay north and south channel samples taken on 6.13.06, 8.4.06, and 6.21.07. North channel samples are indicated as 6.13.06 (●), 8.4.06 (●), and 6.21.07 (●). South channel samples are indicated as 6.13.06 (♦), 8.4.06 (▲), and 6.21.07 (●). A.) AAP cells/ml. Orange line is the no template control. Standard deviations are +/- one deviation B.) % AAP/Bacteria.
Figure 6: Relationships between AAPs and *Synechococcus* (A), Total bacteria (B), and Total chlorophyll (C). Relationships were calculated based on the whole dataset but data here is grouped into locations: Aniwa (●), Futuna (■), Lord Howe (▲), California (○), Molokai (▲), Oahu (●), Kaneohe Bay 6.13.06 (●), 8.4.06 (■) 6.21.07 (▲)
Table 1: List of sampled locations, dates sampled, and the length of the transects conducted. Filter sizes used for size fractionated chlorophyll data are listed. Kaneohe Bay was only sampled once on 6.21.07 with the 5 μm size.

<table>
<thead>
<tr>
<th>Location</th>
<th>Sample Date (mm/dd/yyyy)</th>
<th>Length of Transect (nm)</th>
<th>Salinity</th>
<th>NO₂, NO₃ Nutrients</th>
<th>PO₄⁺ Nutrients</th>
<th>Chlorophyll Filter Sizes (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>07/17/2007-08/02/2007</td>
<td>332.48</td>
<td>x</td>
<td>n/a</td>
<td>n/a</td>
<td>0.22, 0.8</td>
</tr>
<tr>
<td>Oahu</td>
<td>11/5/2006</td>
<td>31.44</td>
<td>x</td>
<td>n/a</td>
<td>x</td>
<td>0.22, 0.8</td>
</tr>
<tr>
<td>Molokai</td>
<td>11/6/2006</td>
<td>31.08</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>0.22, 0.8</td>
</tr>
<tr>
<td>Aniwa</td>
<td>1/21/2007</td>
<td>3</td>
<td>n/a</td>
<td>n/a</td>
<td>x</td>
<td>0.22, 2, 5</td>
</tr>
<tr>
<td>Futuna</td>
<td>1/21/2007</td>
<td>1.59</td>
<td>n/a</td>
<td>n/a</td>
<td>x</td>
<td>0.22, 2, 5</td>
</tr>
<tr>
<td>Lord Howe</td>
<td>2/6/2007</td>
<td>15.79</td>
<td>n/a</td>
<td>x</td>
<td>x</td>
<td>0.22, 2, 5</td>
</tr>
<tr>
<td>Kaneohe Bay</td>
<td>06/13/2006,08/04/2006,06/21/2007</td>
<td>3.7,3.4</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>0.22, 2, 5</td>
</tr>
</tbody>
</table>
Table 2: Regressions analysis of AAPs. n= number of samples, r = r value, p = p-value, m = slope of line, and lastly spearman’s rank value determined with minitab.

<table>
<thead>
<tr>
<th>Variables</th>
<th>n</th>
<th>r</th>
<th>p</th>
<th>m</th>
<th>spearman’s rank value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synechococcus (cells/ml)</td>
<td>94</td>
<td>0.871</td>
<td>&lt;0.005</td>
<td>0.1859</td>
<td>0.832</td>
</tr>
<tr>
<td>Total chlorophyll a (μg/L)</td>
<td>94</td>
<td>0.802</td>
<td>&lt;0.005</td>
<td>3.53E+04</td>
<td>0.852</td>
</tr>
<tr>
<td>Total bacteria (cells/ml)</td>
<td>94</td>
<td>0.784</td>
<td>&lt;0.005</td>
<td>0.0365</td>
<td>0.823</td>
</tr>
<tr>
<td>0-2 um chlorophyll a (μg/L)</td>
<td>77</td>
<td>0.743</td>
<td>&lt;0.005</td>
<td>6.08E+04</td>
<td>0.758</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>94</td>
<td>0.344</td>
<td>&lt;0.005</td>
<td>7.25E+03</td>
<td>0.335</td>
</tr>
<tr>
<td>Nitrate (μM)</td>
<td>25</td>
<td>0.03</td>
<td>0.964</td>
<td>4.98E+04</td>
<td>0.238</td>
</tr>
<tr>
<td>Salinity</td>
<td>72</td>
<td>-0.101</td>
<td>0.399</td>
<td>-1.41E+04</td>
<td>0.424</td>
</tr>
<tr>
<td>Nautical miles</td>
<td>79</td>
<td>-0.181</td>
<td>0.111</td>
<td>-268.27</td>
<td>-0.334</td>
</tr>
</tbody>
</table>
Chapter 3: Aerobic Anoxygenic Phototrophic Bacterial Diversity

Photosynthetic bacteria are found within 5 bacterial phyla (Cyanobacteria, Proteobacteria, Chlorobi, Chloroflexi, and Firmicutes); aerobic anoxygenic phototrophic bacteria (AAPs) cluster within the proteobacteria phyla (Bryant and Frigaard, 2006). Earlier diversity studies of AAPs found them all clustered within the α proteobacteria subclass, closely aligning with their anaerobic counterparts and other chemotaxic bacteria (Yurkov and Beatty, 1998b). However, recent evidence from culture and culture-independent studies have shown AAPs to be more broadly distributed, with members in the γ and β proteobacterial subclasses (Beja et al., 2002). Based on 16S rRNA sequences, AAP bacteria cluster with Roseobacter (α-3), Rhodobacter (α-1), Erythrobacter (α-4), OM60 (γ), and NOR5 (γ) genera. Phylogenetic relationships among 16S and photosynthetic genes are incongruent suggesting that lateral transfer of the photosynthetic superoperon has produced a large diversity of organisms with the same photosynthetic system and pigments, in spite of more conserved photosynthetic genes (Nagashima et al., 1997, Yurkov and Beatty, 1998). The photosynthetic superoperon is commonly used genetically to describe diversity of AAPs due to the conserved regions and the difficulty to target AAPs diverse sequences using 16S rDNA.

Our understanding of how these different AAP groups are distributed through the ocean has been changing drastically as more AAPs are found in different proteobacteria subclasses. Originally the majority of AAPs were thought to be related to α proteobacteria including Erythrobacter strains, but in recent genetic studies most of the diversity is found to fall outside of this group (Beja et al., 2002). Further advances in primer design have revealed γ and α proteobacterial groups dominate the AAP
community (Yutin et al., 2005). For example, in the Baltic Sea AAP γ proteobacteria were detected with FISH probes and dominated the AAP community (Masin et al., 2006).

In the global ocean sampling (GOS) expedition the Roseobacter-like group dominated the oligotrophic AAP community. In addition, the GOS revealed new unknown groups throughout the Atlantic and Pacific Ocean previously missed by current primers (Yutin et al., 2007). Thus, AAPs are far more diverse than originally thought.

Evidence from GOS and other studies have suggested that this diversity is structured along near shore/offshore gradients. For example, coastal waters sampled by the GOS expedition demonstrated AAP groups, like β-proteobacteria, were consistently detected near freshwater sources, as well as other groups that seemed more abundant in coastal waters. The GOS dataset for AAP is the first evidence of coastal and open-ocean types and further research is still needed on a smaller spatial scale to see if these patterns are general and if so, what environmental variables are driving them. In support of the patterns found by the GOS sampling, another global survey reported a strong relationship between AAP diversity and chlorophyll concentrations (Jiao et al., 2007). High diversity was reported when chlorophyll concentrations were the lowest and vise versa for high chlorophyll concentrations. Thus, there is evidence that diversity is structured with respect to environmental gradients and that specific clades maybe ecotypes, are associated with particular ecologies.

While both of these global scale surveys contribute to the understanding of AAP ecology on very large scales, my study bridges the gap between relating the abundance of AAPs to environmental variables and linking these patterns to changes in the phylogenetic diversity. Using natural near shore and offshore environmental gradients
near Pacific Islands, I explored the diversity of AAPs as distance from shore increased using dissociations curves and clone libraries. Abundance and environmental data described in chapter 2 were used to understand how these variables influence AAPs diversity. I hypothesize there will be broad diversity changes between near shore and offshore sites and AAP genetic types will be structured along environmental gradients.

METHODS

Sampling

Eight near shore/offshore transects were conducted within the Pacific Ocean at the following locations: California (28° 36’N, 125° 4’W), Aniwa (18° 10’S, 169° 15’E), Lord Howe (31° 33’S, 159° 05’E), Futuna (19° 32’S, 170° 13’E), Oahu (21° 28’N, 157° 59’W), Molokai (21° 08’N, 157° 02’W), and 2 within Kaneohe Bay located on the island of Oahu (21° 27’ 35”N, 157° 48’ 15” W). Surface water samples were collected in dark amber bottles for environmental DNA. Two of the offshore transects which are located within Kaneohe Bay were sampled three times (6.13.06, 8.4.06, & 6.21.07) (Figure 1).

DNA was collected by filtering 100 or 50 ml of water through 0.22 µm polycarbonate filters (GE Water and Process Technologies) under low vacuum pressure followed by 3 ml of preservation solution (10mM Tris, 100mM EDTA, 0.5 NaCl at a pH of 8.0). Filters were placed in bead beating tubes and stored in a -80 °C freezer. For details on DNA extraction and QPCR techniques see chapter 2 methods.
**Preliminary AAP Diversity Study at Kaneohe Bay**

Due to genetic variability in AAPs, it is important to use primers that capture their full diversity. To assess if I chose primers that would adequately represent the diversity in the Pacific islands, three preliminary clone libraries were performed from 3 stations in Kaneohe Bay (21° 27’ 35”N, 157° 48’ 15” W). Partial *pufM* products were amplified using the primers *pufM*(1669)F (5’GGN AAY YTNY TAY AAY CCN TTY CA 3’) and *pufM*(1840)R (5’CCA TSG TCC AGK GKCT AGA A 3’) and the following PCR conditions: 95 °C for 10 min, then 40 cycles of 95 °C for 1 min, 48.6 °C for 1 min, 72 °C for 5 seconds, and 10 min at 72 °C (modified from Waidner and Kirchman, 2005).

Twenty five microliter reactions were performed using the following solutions: 17 µl of water, 2.5 µl of 10X buffer (Fermentas), 2 µl of 25 mM MgCl₂ (Fermentas), 0.5 µl of 10 mM dNTP’s (Promega), 0.5 µl of 25 µM/L stock for *pufM*(1669)F and *pufM*(1840)R primers, 0.1 µl of Hot Start Taq (Fermentas), and 1 µl of DNA template. For each library, triplicate PCR reactions were run and pooled together when visualizing on a 1% agarose gel stained with ethidium bromide. PCR products were extracted from the gel using a Qiagen gel extraction kit via manufacture’s instructions. Once re-suspended, DNA was inserted into a pCR4-TOPO (Invitrogen) vector and then transformed into E-coli using a TOPO Cloning kit. Clones were grown overnight on LB plates with 50 µg/L of kanamycin (Shelton Scientific) and randomly picked for sequencing. Sequences were cleaned and oriented using Sequencher and aligned in ARB (Ludwig, W. et al., 2004). Neighbor-joining trees were made in ARB.
**PufM Clone Libraries**

Six clone libraries were created to explore the difference between offshore and near shore pufM containing bacterial communities. The criteria for selecting samples to perform clone libraries were to have near shore and offshore stations that had distinct dissociation curves (Figure 7). Two clone libraries each were constructed from the islands of Aniwa, Oahu, and the coast of California. To determine how many clones would be needed to see a change in diversity structure, sequences obtained from Kaneohe Bay preliminary clone libraries were used to create rarefraction curves using DOTUR (Schloss and Handelsman, 2005). Based on those preliminary rarefractions curves, 48 clones per library were chosen for each of the new libraries. Six libraries were constructed as described above, except different primers and PCR program were used to amplify the partial pufM PCR product. PCR primers differed from the Kaneohe Bay PCR reactions in that pufM557F and pufMWAWR were used in the QPCR reaction to determine AAP abundance. The new primers amplify a smaller portion of the pufM gene but are inside of the amplified product used for the Kaneohe Bay libraries. Partial pufM PCR product (240bp) was amplified using pufM557F (5'TAC GGS AAC CTG TWC TAC 3') and pufM_WAWR (5'GCR AAC CAC CAN GCC CA 3') primers and the following PCR conditions: 95 °C for 10:00 min, then 40 cycles of 95 °C for 15 sec, 56 °C for 45 sec, 72 °C for 45 sec. Twenty five microliter reactions were conducted in triplicate with the following reagents (17.77 µl water, 2.5 µl 10X buffer with 15mM MgCl2 (Sigma), 0.24 µl of 10 µM primer stock, 0.25 µl of Jump Start Taq (Sigma), and 2.5 µl of DNA template).
Phylogenetic clusters observed in the clone libraries results were related back to original dissociation curves to understand which clusters represented observed dissociation peaks. Dissociation curves were created for 10 clones from each library. The clones picked were representatives of sequences from major clusters.

**Comparing Clone Libraries using SONS and confirming clone library saturation**

To determine the level of sampling coverage, rarefraction curves were constructed for each library following Schloss and Handelsman (2005). Near shore and offshore libraries at each location were compared using SONS (Schloss and Handelsman, 2006). Near shore and offshore libraries were compared within an island, and were also pooled among all islands and compared. Comparisons within pooled near shore or offshore sequences confirmed that community structure was comparable. Kaneohe Bay was excluded from this analysis due to a lack of an offshore library.

**Examining Diversity of AAPs using Dissociation Curves**

Dissociation curves were measured on pufM amplicons obtained from the AAP enumeration protocol (as described in Chapter 1 methods) using the following protocol: 95 °C for 15 sec, 56 °C for 30 sec, then 95 °C for 30 sec. Between 56 °C and 95 °C the heating rate was at a constant 0.2 °C per second. The change in fluorescence with change in temperature was calculated using the 7300 SDS software and plotted versus temperature. Raw data of the absolute change in fluorescence with change in temperature was imported into Excel for further analysis. Each station’s dissociation curves was
averaged and normalized to the change in fluorescence with change in temperature over proportional abundance. Normalized data was plotted against temperature using Matlab.

Dissociation curves were composed of multiple temperature peaks per sample and to quantify the number of peaks a Matlab code created by Kevin Bartlett was used. The program identified peaks when the slope of the line changed signs. After applying the matlab code, major temperature peaks were manually identified and used for diversity analysis.

RESULTS

*Kaneohe Bay Preliminary Diversity Study Results:*

For Kaneohe Bay 89 sequences were blasted in the NCBI website and all of the top scores were similar to *pufM* sequences with e-values ranging from 2e^{-17} to 9e^{-87}. Only 2 of the Kaneohe Bay sequences were similar to cultured AAP *pufM* sequences. Phylogenetic analysis of Kaneohe Bay sequences revealed a diverse assemblage of sequences with two main clusters containing the majority of the sequences (22, 37 sequences) (Data not shown). Sequences did not cluster with known cultured AAPs but instead clustered with sequences obtained from metagenomics, BAC’s, and clone libraries analysis.
Clone Libraries Results:

Clone libraries were selected based on dissociation curves that would have broad diversity difference based on the concept that different dissociation temperature peaks represent sequences with different base pair combinations for a targeted gene. Aniwa, California, and Oahu’s near shore and offshore dissociation curves were composed of varying number of temperature peaks, as well as different temperatures where the peaks occurred (Figure 7). For example, Aniwa’s near shore (AN1) dissociation curve had one peak at 86 °C. Whereas Aniwa’s offshore (AN6) dissociation curve had three peaks at 83, 85, 88.5 °C. Thus, suggesting one AAP genetic population should be present for the near shore (AN1) clone library and three AAP genetic populations for the offshore library (AN6). California’s near shore (CA2) dissociation curve has three temperature peaks at 82, 84, and 86 °C, whereas the offshore (CA5) curve has one temperature peak at 83°C. Oahu’s near shore and offshore dissociation curves overlapped significantly suggesting similar AAP genetic populations will be found for both near shore and offshore clone libraries.

All sequences obtained from clone libraries were blasted in the NCBI database and the majority of the top scores were identified as pufM genes with e-values ranging from 2 e\(^{-17}\) to e\(^{-11}\). Only California (total sequences 89) and Oahu (total sequences 92) had one sequence that did not have a top match with the pufM gene and these were removed from the analysis. A majority of sequences were identified with pufM sequences obtained from clone libraries, BAC’s, or metagenomics. Aniwa sequences had the largest number of sequences (32%) similar to pufM sequences derived from cultures within the α proteobacteria subclass such as Rhodobacter (α-1), Roseobacter (α-3), and Erythrobacter
For Oahu and California less than 5% of all sequences matched with culture *pufM* sequences, respectively.

High diversity was observed for all libraries and distances between clusters ranged from 94% to ~80% similar. In general within an island, clone libraries grouped into phylogenetic clusters that were dominated by either near shore or offshore sequences. As well as the differentiation between near shore and offshore sequences occurring at the subclass level. For example, California and Aniwa had distinct clusters that were dominated by either near shore or offshore sequences (Figure 8A,B). Oahu’s tree formed 5 distinct clusters each mixed with near shore and offshore sequences within a cluster (Figure 8C).

California’s neighbor-joining tree for CA2 and CA5 had clusters separated by genetic distances of up to 85% similar; these distances suggest genetic difference of at least class or subclass levels (Figure 8A). Near shore sequences (CA2, red in Figure 8A) arranged into five clusters, with three of the clusters containing only near shore sequences. Offshore sequences (CA5, blue in Figure 8B) arranged into four clusters, with two of the clusters containing only offshore sequences. While the other CA5 clusters were dominated by offshore sequences.

Aniwa’s neighbor-joining tree was segregated between AN1 and AN6 sequences. These clusters were genetically separated by a distance of 0.1 to 0.15 substitutions (Figure 8B). Near shore sequences (AN1, red in Figure 8B) arranged into three clusters with two of the clusters dominated by near shore sequences. Offshore sequences (AN6, blue in Figure 8B) were represented by one major cluster, which contained approximately
80% of all offshore sequences. The offshore cluster could be bifurcated into two clusters with a genetic distance of 0.1 substitutions separating them.

Oahu’s neighbor-joining tree had overlapping sequences from both stations within clusters. Five major clusters composed Oahu’s neighbor joining tree and four of these clusters were composed of near shore and offshore sequences. Only one cluster was composed of near shore sequences (OA1, red in Figure 8C). Genetic distances between clusters were generally 0.08 to 0.20 substitutions (Figure 8C).

**Dissociation Curves of Selected Clone Library Sequences:**

Dissociation curves produced from QPCR of *pufM* clones were used to confirm which clusters from the phylogenetic trees were contributing to the original dissociation curve temperature peaks (Figure 7). Almost all of the clusters had distinct temperature ranges as reflected by the original dissociation curves (Figure 7, Figure 8). In general, a majority of clusters could be related back to the original dissociation peaks, while a few clusters had dissociation temperature peaks that were not represented in the original dissociation curve, or vise versa. Therefore, more clones would need to be processed to resolve these differences.

California near shore station (CA2) had four phylogenetic clusters with dissociation temperatures ranging from 84.5 to 86.3 °C. This is consistent with the major peak in the original dissociation curve at 85. °C (Figure 7B, Figure 8A). Two near shore peaks (82, 83°C) from the original dissociation curve were not identified by *pufM* clones. California offshore station (CA5) only had one phylogenetic cluster with temperatures
representing the dissociation peak of 83.3 °C. While two offshore clusters had temperatures ranging from 84.5 to 87.8 °C, which was not identified as peaks.

Aniwa’s offshore station (AN6) had two dissociation peaks at 84.8 and 88.4 °C and appropriate temperature ranges were identified to belong to two different clusters (Figure 7A, Figure 8B). No offshore sequences had a dissociation temperature peak that matched the peak of 83 °C from the original dissociation curve. Aniwa near shore station (AN1) had no partial \( pufM \) clones that had dissociation temperatures that matched the original dissociation peak at 86.1 °C. Higher or lower dissociation temperatures were observed for four main clusters that were composed of near shore sequences. One cluster had dissociation temperatures ranging between 84.1 to 85.5 °C and the other three clusters had higher dissociation temperatures ranging from 87.4 to 88.4 °C.

For Oahu’s near shore and offshore stations, clusters represented dissociation peaks and had distinct temperature ranges. Four clusters were identified to have dissociation temperature peaks of 86-87 °C, while one cluster represented the lower dissociation temperature peak (84 °C). The near shore station (OA1) lacks a cluster dominated by near shore sequences, which would have represented the original dissociation temperature peak of 88.5 °C.

In general, multiple clusters for each station had temperature peaks that represented one major peak in the original dissociation curve. While individual AAP clusters, at a particular genetic distance, should represent a single dissociation peaks it is possible that these different clusters have the same dissociation temperature due to various combinations of genetic deletions and insertions. As well as some genetic distances between clusters are small, therefore small changes in genetic sequences mean
similar dissociation temperatures. Thus genetic identification of the various clusters is needed to determine at what genetic level dissociation curves identify differences between AAP populations.

**Clone Library Community Structure**

Aniwa was the most diverse AAP community based on the high number of OTUs, at the 80% similarity level (Table 3). Aniwa libraries had different community structures and very little overlap occurred between OTUs. In contrast, 43% of Oahu’s OTU’s were shared between near shore and offshore sequences. As well as 75% of the community structure was similar for Oahu’s near shore and offshore communities, at the 80% similarity level (Figure 8C). California had the lowest diversity compared to the rest of the locations, based on the number of OTU’s. Similar to Oahu, California had the same percentage of shared OTU’s, but the similarity between near shore and offshore communities was 52%, at the 80% similarity level.

At the 80% similarity level, all rarefraction curves appeared to be approaching saturation, except for Aniwa station AN1 (Figure 9B). AN1 lack of saturation is explained by the higher number of OTU’s observed.

**Nearshore and Offshore Diversity and Community Structure:**

A global dataset exploring near shore and offshore diversity among all stations was made by combining all of the sequences obtained from Aniwa, Oahu, and California clone libraries. A neighbor joining tree rooted to *Rhodobacter veldkampii* was created to explore if a separation between near shore and offshore sequences observed within a
location would occur when all locations were pooled together (Figure 10). A bifurcation separates a majority of near shore and offshore sequences at a genetic distance of 0.17 substitutions.

To further explore the separation of near shore and offshore sequences, I statistically compared the similarities between near shore/offshore community structures (Table 3). Prior to this comparison, I determined if the combined libraries of near shore or offshore had similar community structure, to then allow for a comparison between near shore and offshore libraries. The near shore library was the most diverse based on number of OTUs, and community structure was only similar between Aniwa and Oahu sequences. California sequences were quite different from the rest of the locations, which influenced the neighbor-joining trees' topology. For the offshore library there were fewer OTUs and the community structure between locations was similar. The variability observed for near shore and offshore libraries community structure suggests that some of the variability is due to location, as well as distance from shore affecting the AAP community composition.

At the 80% sequence similarity level, the composition of the combined near shore and offshore libraries were only were 52% similar. Therefore, 50% of the AAP community did not overlap and was unique to either near shore or offshore sequences. Rarefraction curves saturated at the 85% similarity level.

Environmental variables from chapter 2 were included to determine if the observed separation of near shore/offshore sequences in the combined library was explained by total chlorophyll, total bacteria, and *Synechococcus* concentrations (Figure 11). Variables from the same location clustered together, suggesting environmental
variables explained the lack of AAP community overlap within a location, as well as between locations. For example, *Synechococcus* concentrations have a near shore/offshore separation, where high *Synechococcus* concentrations relate to near shore AAP diversity (top of circle tree) (Figure 11C). Whereas, chlorophyll concentrations cluster with AAP sequences associated with a specific locations and environment (Figure 11A). The patterns of AAP types clustering with environmental variables suggest AAPs diversity could be structured along environmental gradients, as well as AAP genetic types maybe associated with particular environments.

**Broad AAP Diversity Determined by Dissociation Curves**

Based on clone library described above, the formation and disappearance of dissociation peaks can be used as tools to screen for broad changes in community structure. Thus, I used all of the dissociation curves created in chapter 2 and analyzed changes in diversity as a function of distance from shore (a proxy for environmental gradients) (Figure 12, Figure 13).

Each station had one to multiple (maximum 3) dissociation temperature peaks. Dissociation curve temperatures were between 81°C to 88°C. Dissociation curves varied within a location and between locations by the number of temperature peaks observed, as well as at which temperature these peaks occurred at. There were two alternate significant trends that were present: 1.) one dominate peak for stations closest to shore with the number of peaks increasing as distance from shore increased; 2.) multiple peaks were observed for stations closest to shore and number of peaks decreased as distance from shore increased; Aniwa, California, and Lord Howe followed the first trend, Oahu
followed the second trend. Futuna’s near shore dissociation curves were inconsistent, but the offshore stations had a distinct signature.

**DISCUSSION**

In this study, AAPs diversity was found to change within an island, as well as between locations. For example, multiple peaks from dissociation curves were observed to change within an island, as well as between locations. Using clone libraries, the observed changes in dissociation temperature peaks between stations were shown to be related to diverse phylogenetic clusters of AAPs. Previous studies using dissociation curves report similar temperature ranges for AAP *pufM* amplicons, but multiple dissociation peaks have not been observed (Schwalbach and Fuhrman, 2005, Du et al., 2006). The multiple peaks I detected maybe unique to these locations, or derived from the selection of primers used in this study.

The changes in AAP diversity, observed with both clone libraries and dissociation curves, occurred at the subclass or genera level, suggesting the differences in community structure between near shore and offshore locations occurred at the α, β, and γ proteobacterial subclasses level. While genetic distances between clone libraries clusters were inferred from genetic distances calculated from the neighbor joining tree, the specific subclasses α, β, and γ proteobacterial were not identified in this study. While α, β, and γ subclass clusters were not identified in this study, Kaneohe Bay’s dissociation curves near freshwater sources had no major changes in AAP community composition compared to other Kaneohe Bay stations, implying no freshwater β proteobacterial
subclass was present near river stations. Within this dataset observed diversity changes were speculated to be between the γ and α proteobacteria subclasses.

Previous studies have demonstrated that γ and α proteobacterial subclasses dominate AAPs diversity (Yutin et al., 2005, Yutin et al., 2007, Masin et al., 2007). Due to the observed separation between near shore and offshore communities it is speculated the changes in community structure was between open ocean and coastal γ and α proteobacteria AAP types. For example, γ proteobacteria have been shown to dominate AAP diversity in most marine environments; this subclass contains two phylotypes obtained from both coastal and open ocean ecosystems (Hu et al., 2006, Cho et al., 2007). For α proteobacteria, Roseobacter and Roseobacter-like sequence types contribute most substantially to oligotrophic environments compared to other α proteobacteria genera (Yutin et al., 2007, Oz et al., 2005). In the GOS expedition, similar dominance of α and γ proteobacteria was observed, as well as AAP genera presence varying depending on the environment (Yutin et al., 2007).

Along these offshore transects environmental variables chlorophyll, total bacteria, and Synechococcus were shown to explain the variability in AAPs’ diversity. Therefore, suggesting AAPs diversity is structured along environmental gradients. In a previous study of chlorophyll concentrations, AAP abundance, and AAP diversity, AAP abundance and chlorophyll concentrations were shown to have a positive relationship, where as AAP diversity and chlorophyll concentrations were negatively correlated (Jiao et al., 2007). This study observed that chlorophyll concentrations grouped with specific AAP phylogenetic clusters, while having a positive relationship with AAP abundance. In contrast to Jiao and collaborators, the relationship between AAP community structure and
chlorophyll concentration varied within a locations as well as between locations and suggests other environmental variables other than chlorophyll influence AAP diversity. For example, Kaneohe Bay’s diversity was consistent along both transects, even though Kaneohe Bay had the largest range of chlorophyll concentrations. Elevated AAP abundance observed at the closest station to the river may infer that AAP abundance and diversity is influence by particles, therefore increasing the number of AAP particle associated species, which are contained within the γ proteobacteria subclass. Whole genome analyses of AAP γ proteobacteria strain Congregibacter litoralis KT71, suggests these organisms prefer suboxic environments with potential enzymes production common to particle attached bacteria (Fuchs et al., 2007). The increased particle load near Kaneohe Bay’s river station suggests a suitable habitat for coastal γ proteobacteria.

For all the other locations AAP abundance and AAP diversity were influenced by location and types of benthic habitat. For example, Aniwa and Lord Howe had one dissociation peak near shore where AAP abundance was high, and diversity became more complex as AAP abundance decreased. As mentioned in Chapter 2 both of these islands had barrier reefs and lagoons. In addition to chlorophyll, clone library results demonstrated a separation of near shore and offshore AAP diversity correlating to Synechococcus concentrations similar to the relationship described with AAP abundance. As in chapter 2, I suggested this relationship is due to a common unidentified environmental variable influencing both populations because they compete for similar resources, or that Synechococcus produces DOM that influences AAP diversity and abundance.
In conclusion, this study reports that AAP are phylogenetically diverse and that some phylogenetic clusters appear to be associated with certain environments. For example, the diversity of AAPs is correlated with chlorophyll, AAP abundance, *Synechococcus*, and total bacterial concentrations. Different AAP diversity types suggest these groups of AAPs perform different ecological roles depending on the environment they occupy. Obtaining some of these sequence types as culture representatives will help investigate how different AAPs might have different amounts of bacteriochlorophyll a pigment, growth rates, or rates of carbon dioxide fixation. These details are imperative to understanding AAPs importance in the oligotrophic and coastal carbon and energy cycles.
Figure 7: Dissociation curves from the Island of Aniwa (A), California (B), and Oahu (C) used for clone library studies. Squares represent peaks in dissociation curves.
Figure 8: Phylogenetic relationships (neighbor-joining) among AAP clones from (A) California, (B) Aniwa, and (C) Oahu inferred from partial *pufM* sequences (240 bp). Pie charts represent the proportion of sequences that contribute to nearshore (red) or offshore (blue). Temperature ranges represent the peaks of the dissociation curves found in those clusters, with the number of observations in parentheses.
Figure 9: Rarefractions curves for Kaneohe Bay station SC2 (A), Aniwa AN1(B), and Oahu OA3(C). Dark blue line 100% similarity line, green line 99 % similar, red line 94% similar, and light blue 80% similar.
Figure 10: Neighbor joining tree of clones inferred from partial pufM (240 bp) sequences from Aniwa, Oahu, and California. Green bars indicate clones from near shore, blue bars indicate clones from offshore.
Figure 11: Relative proportion of environmental variables total chlorophyll (A), total bacteria (B), or Synechococcus (C) overlaid onto a neighbor joining tree of clones inferred from partial *pufM* sequences from Aniwa, Oahu, and California. Blue branches indicate offshore sequences, green branches indicate near shore sequences.
Figure 12: Dissociation curves for Aniwa, Futuna, Kaneohe Bay southern channel (KB-SC) and northern channel (KB-NC) separated into categories of distance from shore (0-0.5, 0.5-1, 1-2, and 2-4 nautical miles). Squares represent temperature peaks.
Figure 13: Dissociation curves for California, Lord Howe, Oahu, and Molokai transects separated into categories of distance from shore (0-5, 5-10, 10-100, 100-300, and > 300 nautical miles). Squares represent temperature peaks.
Table 3: Comparing onshore and offshore clone libraries community structure for California, Aniwa, and Oahu. Number of OTU's were determined using DOTUR for both stations. Number of sequences, OTU’s shared/Total OTU’s, and Community Structure were calculated using SONS. Community Structure Similarity is calculated based on overlapping OTUs, as well as the relative abundances among the shared OTUs between the communities.

<table>
<thead>
<tr>
<th>Locations</th>
<th>Distance level</th>
<th># of sequences</th>
<th># of OTU's</th>
<th>OTU’s shared/Total OTU’s</th>
<th>Community Structure Similarity</th>
<th>Community Structure Similarity SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>78%</td>
<td>88</td>
<td>7,7</td>
<td>0.43</td>
<td>0.52</td>
<td>0.13</td>
</tr>
<tr>
<td>Oahu</td>
<td>78%</td>
<td>91</td>
<td>7,7</td>
<td>0.43</td>
<td>0.75</td>
<td>0.12</td>
</tr>
<tr>
<td>Aniwa</td>
<td>80%</td>
<td>93</td>
<td>14,14</td>
<td>0.14</td>
<td>0.042</td>
<td>0.03</td>
</tr>
<tr>
<td>All</td>
<td>80%</td>
<td>272</td>
<td>22,22</td>
<td>0.41</td>
<td>0.52</td>
<td>0.083</td>
</tr>
</tbody>
</table>

Nearshore

- CA vs. OA 80% 91 19,19 0.13 0.0059 0.0049
- CA vs. AN 80% 93 19,19 0.11 0.059 0.038
- AN vs. OA 80% 92 19,19 0.3 0.21 0.076

Offshore

- CA vs. OA 80% 90 9,9 0.25 0.3 0.095
- CA vs. AN 80% 88 9,9 0.17 0.99 0.0051
- AN vs. OA 80% 84 9,9 0.29 0.32 0.01
References:


45. Palenik, B., Ren, Q., Dupont, C., Myers, G., Heidelberg, J., Badger, J., Madupu, R., Nelson, W., Brinkac, L., Dodson, R., Durkin, A., Daugherty, S., Sullivan, S.,


