THE BIOGEOGRAPHY OF CENTRAL PACIFIC CORAL REEF FISHES

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI`I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ZOOLOGY (ECOLOGY, EVOLUTION AND CONSERVATION BIOLOGY)

MAY 2012

By

Craig A. Musburger

Dissertation Committee:

Kim Holland, Chairperson
Brian Bowen
Ben Brooks
John Stimson
Andy Taylor

ACKNOWLEDGMENTS

This project would not have been possible without the help of many people and organizations. I would like to especially acknowledge the following people whose help was invaluable: Jim Maragos, Rusty Brainard, Ed DeMartini, Bob Schroeder, Karen Koltes, Sylvia Pinca, Maria Beger, Randy Kosaki, Peter Houk, Eric Verheij, Yim Goulong, Larry Basch, David Obura, Elizabeth Stephenson, Bruce Mundy, Jack Randall, Joe Laughlin, Mark Readie, Paula Ayotte, Brett Schumacher, Kelly Boyle, Ariel Rivera, Sarah McTee, Darla White, Jeff Eble, Michelle Gaither, Peter Vroom, Brian Greene, Berry Muller, Melba White, and James Matayoshi, as well as the following institutions and organizations: NOAA CRED, USFWS, TNC, American Samoa DMWR, US Department of Interior – Insular Affairs, Palau Conservation Society, Marshall Islands Marine Resource Agency, NRAS, Rongelap Atoll Local Government, and the New England Aquarium.

ABSTRACT

The distribution and abundance of coral reef fishes is undoubtedly affected by a variety of factors. While much previous research into dispersal success of reef fishes has focused on species' ability to survive for long periods in the pelagic larval stage, this alone cannot explain biogeographic patterns. Consequently, a novel approach was developed to examine another potentially important factor contributing to dispersal ability: the number of offspring propagules produced at a source location. A large scale data collection effort was undertaken around islands of the Central and Western Pacific to facilitate testing of this hypothesis. Since direct measurement of larval output on this scale is not possible, a proxy for larval output – adult population density – was used as the predictive variable in models used to identify factors underlying dispersal patterns. *In situ* visual sampling was conducted over seven years on 38 islands across the Pacific. Over 750 species and 300,000 individual fish were identified and sampled along 700 transects. Analysis was focused on the large biogeographic barrier separating Johnston Atoll and the Hawaiian Archipelago from the rest of the study islands. The regions most likely to be sources of fish larvae reaching the Johnston-Hawaii system were sampled using robust, standardized, quantitative fish census methods. Methodological precautions were taken to ensure that the types of population estimates used in the analyses were robust in space and over time. Novel statistical methods were developed to evaluate the strength of ecological data and two types were identified as suitably robust and reliable; numeric density from belt-transect surveys and presence-absence records from large search area roving diver surveys. Biomass estimates and presence-absence data from belt transects

were deemed inadequate or inappropriate for these analyses. Highly significant results of regression models confirmed a relationship between adult population density and long distance dispersal in 3 of 5 fish families included in the analysis. This finding is novel and offers a significant new approach to the study of the biogeography of marine species.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:
INTRODUCTION

Biogeography is broadly defined as the study of the distribution of living things. Scientists have been interested in this field since the 19th century, and its creation is often credited to the work of Alexander von Humboldt who extensively documented the botany of Latin America in his landmark publication, Kosmos (Humboldt 1867). Since the time of Humboldt, the major works of Alfred Wallace (Wallace 1906) and Charles Darwin (Darwin 1872) built upon this science, using it to support the major scientific development of the 19th century - the theory of evolution. By understanding the relationships between living things and their spatial distributions, the most fundamental paradigm shift in the biological sciences first became clear to modern biologists.

The field has progressed since its early days as approaches have become more quantitative, explorations have become more extensive, and new theories have been proposed, tested and reformed. Yet, while progress has been made, the knowledge of the distribution of living things is far from complete, and scientists have continued to investigate relationships between species and space in hopes of better understanding the forces that have made the planet what it is today and what it may be in the future.

Insular areas provide a model system for the study of biogeography. With easily identifiable boundaries, islands provide ideal study conditions where hypotheses are more easily tested. This fact was not lost on the fathers of modern biogeography, Robert MacArthur and E.O. Wilson. Their publication of *The Theory of Island Biogeography* (MacArthur and Wilson 1967) was one of the major contributions to the modern view of biogeography as a science, and it sparked a renewed interest in this discipline. While it initially attempted to describe patterns of distribution of species on real "islands," the theories presented in this book were quickly extended to apply to patches of land area that could be considered to be *like* islands when barriers to species ranges were able to be identified. The work done by MacArthur, Wilson, Simberloff, and others (Simberloff and Wilson 1969; Simberloff and Wilson 1970) impacted the study of genetics, conservation biology (Simberloff and Abele 1976), and ecology (Simberloff 1974). Their work provided quantitative predictive models that could be tested in the real world, and it inspired countless subsequent studies. Incorporating terms like extinction rates, immigration and emigration rates, degree of isolation, and island size, quantitative models were formed which could be used to predict things like species richness. One of their important contributions to the study of biogeography was the introduction of the species-area relationship, which relates the species richness of an island to its size.

A subsequent important development in the study of species distributions came in 2001 with the publication of Stephen Hubbell's *The Unified Neutral Theory of Biodiversity and Biogeography* (Hubbell 2001) which attempted to further explain patterns of species richness and relative abundance. Hubbell's theory assumed that all species within a given trophic group are equally likely to dominate their particular local area and that their relative abundances will change randomly through time, but that at any given time, one or a few species will dominate that part of the ecological community. While tests of Hubbell's hypothesis have offered mixed results, the fact that he again contributed to the discussion of biogeography through the use of a readily tested quantitative model shows the maturity of the field as it has progressed from a largely descriptive science into one that focuses more on hypothesis driven quantitative investigations.

What is a species?

The commonly accepted definition of a species is a group of individuals that can reproduce and whose offspring are fertile. The creation of new species is a topic of much study, and there are two commonly accepted notions of how new species can arise. The one uniting factor that must be in place for a new species to arise involves some type of genetic or geographical isolation (Mayr 1942). A parent population must be impacted by some type of isolation into two separate populations for there to be incorporation of genetic differences between the two populations, and this separation must occur for a long enough time period for the genetic differences between the two populations to develop to such an extent that the two new populations – should they become reunited –

could not produce viable offspring. This isolation could occur within the same geographic area (sympatric) or in two or more areas (allopatric). The commonly used model to describe the potential for sympatric speciation would be disruptive selection where two different genotypes within a single population showed increased fitness over a heterogenous genotype. For example, if two feeding habits among a bird species exist which include eating fruits from the tops and the bottoms of the same tree. Over time, if the group feeding only on upper fruits saw a fitness increase with genotype *AA* and the other group of birds feeding exclusively on the bottom fruits saw a fitness increase with genotype *aa*, the populations could over time become isolated to the point that they did not interbreed, and eventually two distinct species could exist (Dieckmann and Doebeli 1999, Smith 1966). Examples of sympatric speciation in the real world are rare and controversial (Via 2001). Allopatric speciation – the much more common and likely scenario - must involve some type of isolating event that physically separates a once continuous population into two or more distinct populations (Mayr 1969). The mechanisms of isolation can be broadly divided into two types – vicariance and dispersal. Vicariance refers to the idea that a population is met with some type of change in the environment that causes the population to be divided into more than one distinct population. An example would be the formation of a rift between continents that separates two land locked populations as the shifting of tectonic plates drives the two continents further apart making intermixing between the two isolated populations impossible. Another example could be the isolation of mountainous areas near the ocean as sea level rise causes once adjoining peaks of the same island to become two separate islands. Dispersal refers to a different type of isolating event where a small number of individuals from a population are able to travel to another location which is adequately isolated from the original location so that gene flow between the two areas is restricted, and genetic changes are incorporated into each of the two new populations at such a rate that eventually interbreeding becomes impossible in a process known as peripatric speciation.

Insular systems like the islands of the central and western Pacific offer a great opportunity to study the dispersal model for speciation. Most of these islands are of volcanic origin and have been geographically distant from one another for the entire duration of their existence. The distribution of species around these areas is often described as a stepping stone type distribution where certain species have been able to make the jump between areas and others have not (MacArthur and Wilson 1967).

While the models of MacArthur and Wilson and Hubbell make predictions about how many species would be likely to inhabit each island and the relative abundances of various species at each island, they do not offer any explanation as to why certain species may become more widely distributed than others or the role of the speciation process.

In marine systems, one aspect of species' life histories has been the focus of most of the research into likelihood of long-distance dispersal: pelagic larval duration (PLD) (Thresher & Brothers 1985; Thresher et al.1989; Wellington & Victor 1989; Bonhomme & Planes 2000; Victor & Wellington 2000; Zapata & Herron 2002; Mora et al 2003). The life cycle of many marine species involves a distinct larval stage which is when most long-distance dispersal occurs. Within coral reef associated fauna there is very little long-range dispersal of adult individuals. The open water between islands is often poor in food, nutrients, and obviously the shelter that these species depend on for their survival. However the larval stage of these species is often well suited for extended periods of time in the pelagic realm (Sale 2002). Species have shown different propensities for the length of time which they are capable of surviving in their larval stage. The dominant theory for previous studies of species distribution likelihood has centered around the idea that the longer a species is able to survive in its larval state, the more likely it is to have a widespread distribution. The idea is simple. If a species is able to remain adrift for a longer period of time, the stochastic forces of wind and currents which largely drive the passive dispersal of these larvae will be more likely to bring a given larva close enough to a potential home so that the larva will be able to settle out and successfully grow into a reproductive adult at the destination location.

Tests of this theory have shown mixed results (Lester and Ruttenberg 2005). While some studies have shown correlations between PLD and species range, these correlations are often very weak if they are detectable at all (Zapata and Herrón 2002, Victor and Wellington 2000). At the very large end of range size, a threshhold PLD of 45 days has been used to suggest a larval duration at which widespread distribution would be expected (Brothers and Thresher 1985). However, within the middle ranges of PLD, the relationship between larval duration and range size or genetic connectivity is not seen (Lester et al. 2007; Weersing and Toonen 2009). Additionally, PLD has been observed to be dynamic even within a single species depending upon time of year and by location where a species is sampled (Bay et al. 2006). Clearly, alternative mechanisms are also at work in determining a species' geographic distribution.

One possible alternative is that the quantity of larval production at a source location may influence likelihood of long distance dispersal. Likelihood of a species' offspring reaching a distant location could be expected to be very low when the distance between sites is far. While a model that considers duration in the water column is logical, the next reasonable idea to test would be this: how would the quantity of larvae produced at the source location affect this likelihood. Surely, it seems reasonable to expect that if a parent population creates a very small number of offspring, the likelihood of dispersal would be lower than if a population produces a huge number of offspring. The mixing that occurs over time during the dispersal process presumably would favor the population which has expelled a larger number of larvae into the mix. While this hypothesis seems reasonable, the practicalities of working in these systems have made testing it impossible to date (Lester and Ruttenberg 2005). It is the principal focus of this PhD dissertation to attempt to test this hypothesis within coral reef fish populations of the Central and Western Pacific Islands. And while direct measurement of the number of eggs and sperm expelled into the water column by a large number of coral reef fish species was not practical or possible, care was taken to find a suitable proxy for this measurement in the form of adult population density.

To test this hypothesis, several steps were undertaken. Each portion of the investigation leading up to the test is presented here as a separate chapter under the following headings:

Chapter 2: Patterns of Distribution of Coral Reef Fish Species in the Central Pacific

Chapter 3: Practicalities and Limitations in Assessing Reef Fish Populations Using Visual Survey Methods

Chapter 4: Reef Fish Population Temporal Stability

Chapter 5: The Relationship between Source Population Abundance and the Likelihood of Long Distance Dispersal in Coral Reef Fishes

CHAPTER 2:
PATTERNS OF DISTRIBUTION OF CORAL REEF FISH SPECIES IN THE CENTRAL PACIFIC

Abstract:

The relationships between the shallow water fish fauna at 38 Central and Western Pacific Islands was investigated. Fish were surveyed using a combination of Belt-Transects and Roving Diver Surveys. Relationships between islands were enumerated using multivariate cluster analysis procedures using the software package Primer-E. Regional relationships largely met with expectations based on geographic proximity and archipelagic boundaries. Main regions delineated included: 1. Johnston Atoll; 2. Northwest Hawaiian Islands; 3. Main Hawaiian Islands; 4. Marshall Islands; 5. Palau; 6. American Samoa; 7. Wake Atoll; 8. Nukunono Atoll; 9. Northern US Line Islands; 10; US Phoenix Islands and Equatorial Line Islands; 11. Swain's Island; and 12. The Southern Phoenix Islands. The major break in the regions occurs between the Johnston Atoll/Hawaiian Islands group and the rest of the regions. This is a well known distinction as the fish fauna of Hawaii is known to be highly differentiated from that of other regions in the Pacific. The identification of the major break between the Johnston Atoll/Hawaiian Islands and the rest of the Central Pacific serves as the foundation for further testing the hypothesis that the abundance of individuals in source populations is a major factor affecting likelihood of long distance dispersal presented in subsequent chapters.

Introduction:

The study of the distribution of coral reef fish species in the Central Pacific is far from complete. The huge number of islands, the long distances between them, and the practical difficulties of surveying coral reef fishes means that even today, new expeditions to many of these areas will inevitably produce range extension records, and even the discovery of previously unknown fish species. This potential for new discovery is even possible in shallow and easily accessible depths. Continued efforts are being conducted to simply define the ranges of many species of fish and to produce complete species records for island locations. While certain parts of the Pacific are very well studied (Hawaii for example), others are less well described. The relationships between the different islands of the Pacific are even less well known. One region which has been largely unstudied is the isolated and remote Central Pacific. This study represents one of the most comprehensive investigations into the patterns of fish species distributions within this area to date.

Methods:

Fish were surveyed at 38 islands (Fig. 2.1 and Table 2.1) by a single observer (the author). Access to the study sites was possible through the generous support of many partner agencies. Extended ship and land based expeditions were undertaken over the course of seven years to collect the data needed for this investigation.

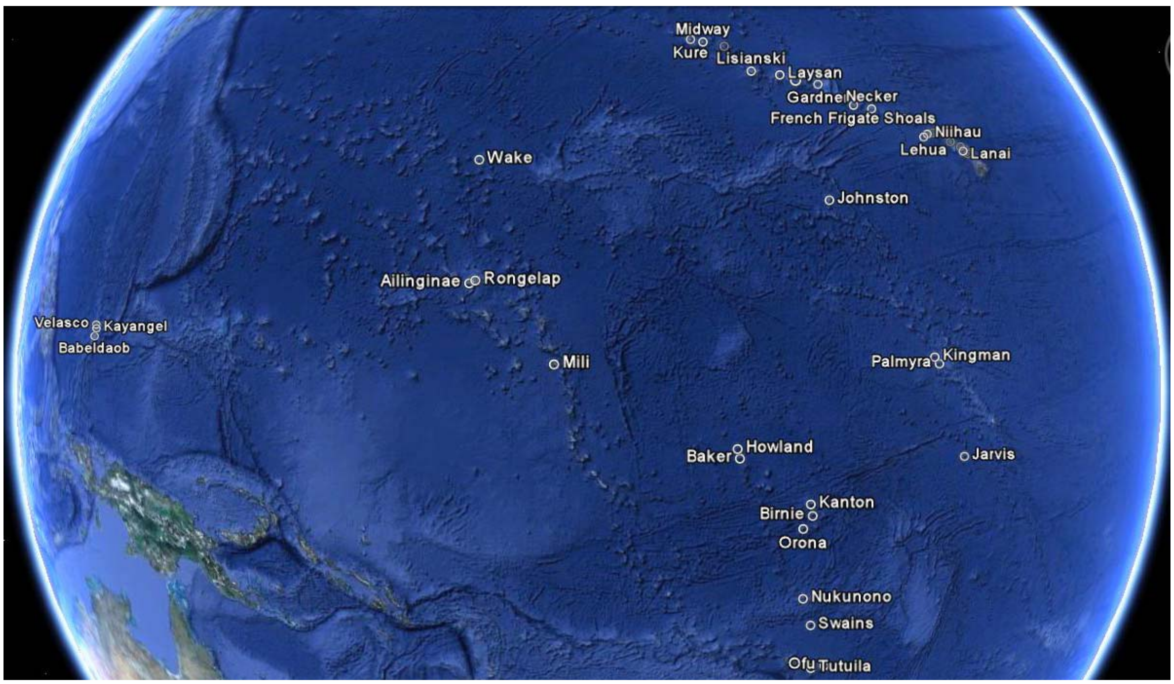

Figure 2.1: Study sites included 38 islands ranging from the nation of Palau in the west to American Samoa in the South and up to the Hawaiian Islands in the North and East. This and subsequent maps created using Google Earth.

At each island, multiple study sites were selected haphazardly while attempting to provide the best possible coverage of all directions of exposure and habitat types. At each site, a survey was conducted using closed circuit SCUBA which limited depths to approximately 30 m. At all sites a roving diver survey was conducted where the observer collected records of all fish species observed at the site during the entire duration of the dive which lasted between 60 and 90 minutes at each site. When fish could not be identified to the species level in situ, video records were collected for subsequent identification. While specimen collection would be a more ideal method in some cases, the protected status of many of the study areas precluded extractive methods from being used in assessing the species composition of the study areas. Since visual identification was conducted primarily in the field, identifications were made based on visual morphology (Myers 1999; Randall 1996; Randall 2002; Randall 2005; Randall 2007). While some genetic analyses have called into question the reliability of morphology as a definitive tool in species classification (Schulz et al. 2006), it is a commonly used approach (Brock 1982) and was the only method available for this research as collecting permits were not available for many of the study locations.

Following collection of species presence/absence records for all islands, multivariate statistical methods were used to determine the relationships between islands. Analyses were conducted using Primer-E software. Species checklists were generated for each island surveyed, and Bray Curtis distance was calculated between each of the island pairs using the equation:

$$BC_{ij} = \sum \frac{|n_{ik} - n_{jk}|}{(n_{ik} + n_{jk})}$$

where *n* is the population density (or in this case the binary response of 1=present and 0=absent) of species *k* at islands *i* and *j*.

The resultant similarity matrix was then processed using the hierarchical cluster analysis procedure in Primer-E using the Group Average cluster mode. The resulting dendrogram was used to evaluate and identify clusters of islands that shared similar fish fauna compositions. Additionally, non-metric multidimensional scaling (nMDS) plots were generated to graphically represent the relationships between the islands.

To test for spatial autocorrelation of the data, a Mantel test was performed on two matrices that contained geographic distance between all study islands and the Raup-Crick dissimilarity index between all study site pairs. The Mantel test and Raup-Crick dissimilarity calculations were performed in R using the ade4 and vegan packages.

Results:

A total of 772 fish species were recorded at 38 islands. Most species were identified in situ by the author, but when positive underwater identification was not possible, photographic and video records were used to identify species in the lab. Over 700 belt-transect surveys were conducted which enumerated over 300,000 individual fish.

Abbreviation	Name	Political Region	Latitude	Longitude
AIL	Ailinginae	Marshall Islands	11° 6'57.53"N	166°23'22.12"E
BAB	Babeldaob	Palau	7°26'45.47"N	134°32'51.93"E
BAK	Baker	US Phoenix Islands	0° 8'41.42"N	176°31'34.99"W
BIR	Birnie	Phoenix Islands	3°41'15.16"S	171°34'9.60"W
FFS	French Frigate Shoals	Northwest Hawaiian Islands	23°43'31.36"N	166°12'14.45"W
GAR	Gardner	Northwest Hawaiian Islands	25°12'19.44"N	168°48'25.61"W
HOW	Howland	US Phoenix Islands	0°44'44.79"N	176°40'51.68"W
JAR	Jarvis	US Line Islands	0°23'24.08"S	160° 1'10.30"W
JON	Johnston Atoll	Johnston Atoll	16°41'8.12"N	169°31'36.25"W
KAN	Kanton	Phoenix Islands	2°56'2.09"S	171°43'53.49"W
KAU	Kauai	Main Hawaiian Islands	22° 2'0.40"N	159°32'32.85"W
KAY	Kayangel	Palau	8° 1'11.31"N	134°42'41.37"E
KIN	Kingman	US Line Islands	6°17'19.38"N	162°27'41.24"W
KUR	Kure	Northwest Hawaiian Islands	28°23'35.07"N	178°18'12.86"W
LAN	Lanai	Main Hawaiian Islands	20°48'16.41"N	156°56'29.01"W
LEH	Lehua	Main Hawaiian Islands	21°38'53.04"N	160°32'38.79"W
LIS	Lisianski	Northwest Hawaiian Islands	25°58'25.69"N	173°58'18.99"W
MAI	Maui	Main Hawaiian Islands	20°43'44.40"N	156°18'43.04"W
MAR	Maro	Northwest Hawaiian Islands	25°25'0.00"N	170°34'60.00"W
MID	Midway	Northwest Hawaiian Islands	28°12'50.41"N	177°22'5.15"W
MIL	Mili	Marshall Islands	6° 3'40.25"N	171°51'13.71"E
MOK	Molokini	Main Hawaiian Islands	20°37'47.06"N	156°29'58.73"W
MOL	Molokai	Main Hawaiian Islands	21° 7'51.32"N	157° 5'1.33"W
NII	Niihau	Main Hawaiian Islands	21°53'56.64"N	160° 7'14.94"W
NUK	Nukunonu	Tokelau	9°13'44.36"S	171°51'41.62"W
OAH	Oahu	Main Hawaiian Islands	21°27'0.36"N	157°59'46.29"W
OFU	Ofu	American Samoa	14°14'58.84"S	169°42'54.51"W
OLO	Olosega	American Samoa	14°11'12.03"S	169°36'53.71"W
ORO	Orona	Phoenix Islands	4°30'54.16"S	172°10'36.49"W
PAL	Palmyra	US Line Islands	5°50'34.88"N	162° 5'58.35"W
PHR	Pearl and Hermes Reef	Northwest Hawaiian Islands	27°55'3.88"N	175°45'51.24"W
RON	Rongelap	Marshall Islands	11°17'28.53"N	166°46'42.53"E
ROS	Rose Atoll	American Samoa	14°33'49.50"S	168°10'0.24"W
SWA	Swains Island	American Samoa	11° 9'33.94"S	171° 7'35.58"W
TAU	Tau	American Samoa	14°15'58.86"S	169°28'10.58"W
TUT	Tutuila	American Samoa	14°20'47.29"S	170°42'11.00"W
VEL	Velasco Reef	Palau	8°17'17.46"N	134°37'14.68"E
WAK	Wake Atoll	Wake Atoll	19°16'5.94"N	166°36'18.48"E

Table 2.1: Islands surveyed for this investigation. Abbreviations are shown to help understand the dendrogram. Political Region does not reflect any faunal relationships.

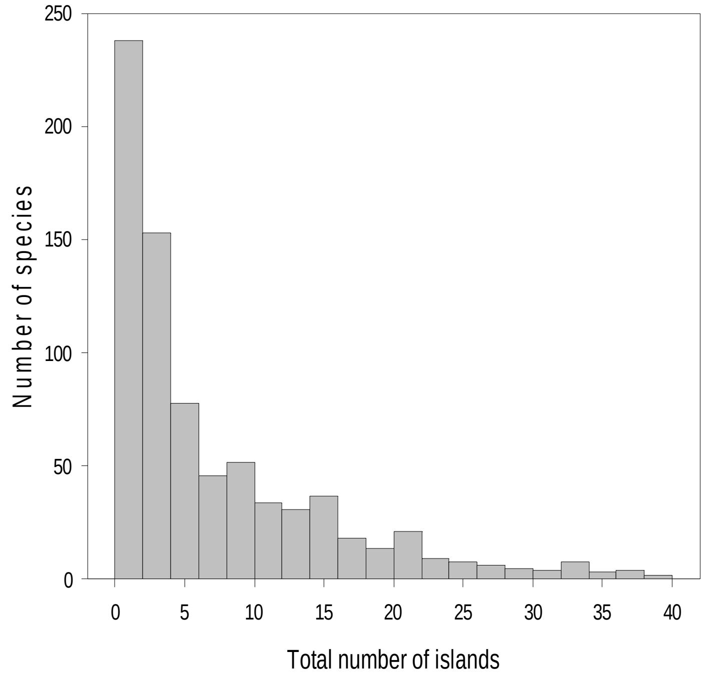

Figure 2.2: Histogram of the total number of islands where each species was found. Single island endemics dominate while relatively few species were widely distributed over the entire study area.

The vast majority of species showed some type of range restrictions and were only recorded on a few islands (Fig. 2.2). A total of 238 species were only recorded on a single island. While endemism on the scale of a single island is very rare in coral reef fishes, this pattern in the data certainly indicates that less than a perfect census was conducted of all species at every island. However, 138 of the 238 species recorded only on a single island were recorded on Babeldaob – the highly diverse single high island surveyed in Palau.

Only two species were recorded on every island surveyed, the arc-eye hawkfish (*Paracirrhites arcatus*) and the multi-banded goatfish (*Parupeneus multifasciatus*). There are many species of coral reef fish that are known to have widespread distribution and which would be expected to occur at all the islands surveyed if a complete census was possible. Time constraints and the realities of underwater visual census methods prohibited this from being possible during this study.

The cluster analysis and resulting dendrogram (Fig. 2.3) elucidated distinct relationships between islands that largely met with expectations based on island proximity and geographic relationships.

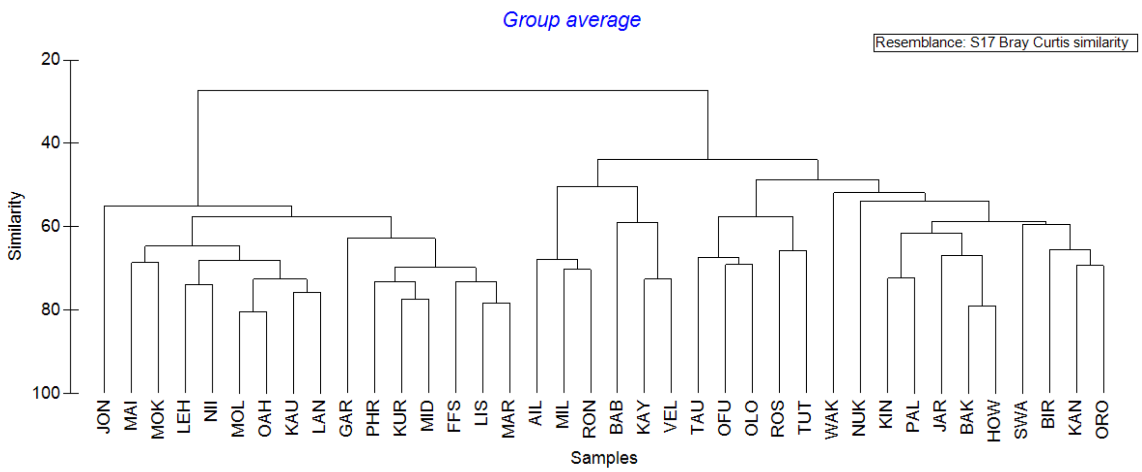

Figure 2.3: Hierarchical cluster analysis dendrogram of the relationships between all study sites based on the presence and absence of observed reef fish species. Analysis performed using Primer-E software and based on Bray-Curtis similarity index and the group average clustering scheme.

The most obvious break in the relationships between island groups separates Johnston Atoll along with all the Hawaiian Islands from all the other islands surveyed. This distinction occurs at the low similarity level of 28 – indicating a major distinction between these two groups of islands. Continuing down the branches of the dendrogram, the next major break isolates the six Micronesian islands in Palau and the Marshall Islands from a cluster that includes Wake Atoll, the equatorial islands, and all the south Pacific Islands surveyed. The fourth major break in the dendrogram removes the islands of American Samoa from this main cluster of equatorial and south Pacific Islands, and then the fifth break isolates Wake Atoll from the remaining Equatorial and South Pacific Islands (Fig 2.4).

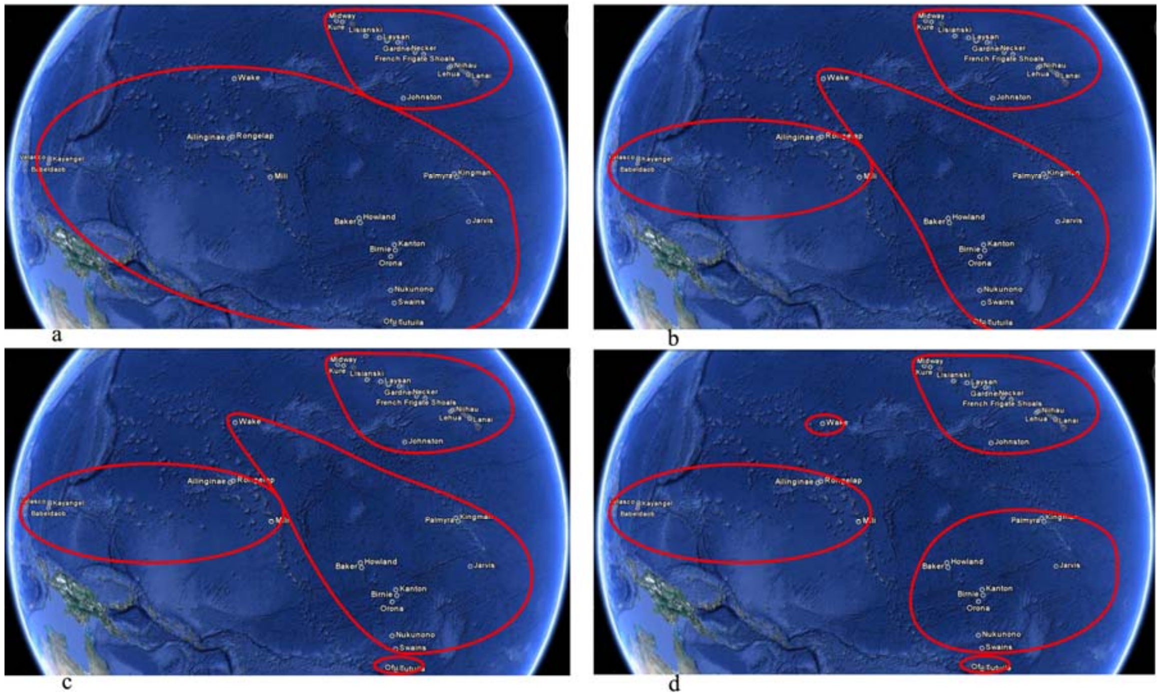

Figure 2.4: Maps showing the first four breaks seen in the cluster analysis. Section (a) reveals the first break isolating the Hawaiian Islands from the rest of the Pacific, while (b), (c), and (d) show subsequent divisions into 3,4, and 5 clusters respectively.

Proceeding down the cladogram to the similarity level of 60, there are thirteen island groups that are identified as distinct clades. They can be described as: 1. Johnston Atoll; 2. Northwest Hawaiian Islands; 3. Main Hawaiian Islands; 4. Marshall Islands; 5. Babeldaob (Palau High Island); 6. Northern Palau (Atoll & Seamount); 7. American Samoa; 8. Wake Atoll; 9. Nukunono Atoll; 10. Northern US Line Islands; 11; US Phoenix Islands and Equatorial Line Islands; 12. Swain's Island; and 13. The Southern Phoenix Islands. As expected, geographic proximity seems to highly influence relationships between island fish fauna. And while it appears that island type also seems to play a part in faunal similarities between islands, further investigation would be required to determine the validity of this relationship as there are few areas in this study where islands of different type are situated close enough together to examine this phenomenon.

As a precaution against potentially overweighting the importance of species that were recorded only on a single island in the cluster analysis procedure, this test was repeated on a subset of the data with all species that were recorded only on a single island removed from the analysis. The results from this test were nearly identical to the test performed on the data with all species included with the only difference being that French Frigate Shoals clustered with the Main Hawaiian Islands as opposed to with the Northwest Hawaiian Islands (Fig. 2.5).

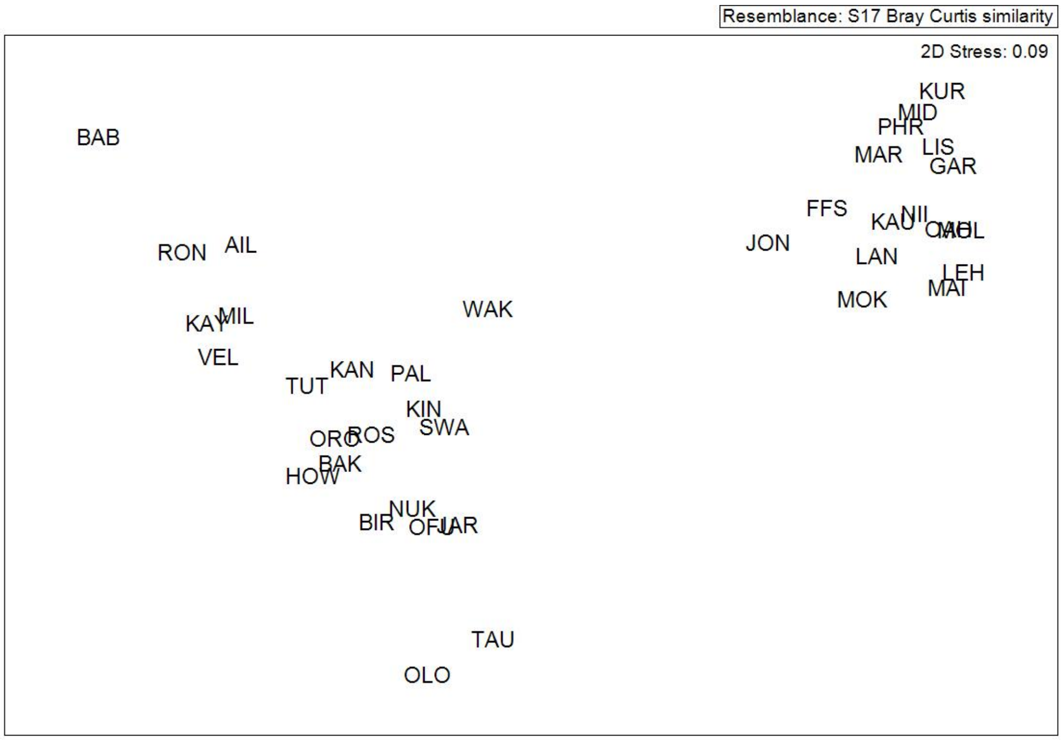

Figure 2.5: A nMDS plot based on the presence and absence of observed reef fish species at all islands in the study.

The nMDS plot based on the analysis of all species recorded in the study (Fig. 2.5) reinforces the main division between the Johnston Atoll/Hawaiian Island fish fauna from the rest of the surveyed islands. The Micronesian islands (Palau and the Marshall Islands) appear closely related, but then the equatorial and South Pacific islands become slightly jumbled in terms of their apparent relationships with islands of Tokelau, Samoa, the Line and Phoenix Islands all intermixed.

A map of the study region with the thirteen clusters highlighted shows clearly how the cluster analysis results fit tightly with the geographic relationships between islands (Fig 2.6).

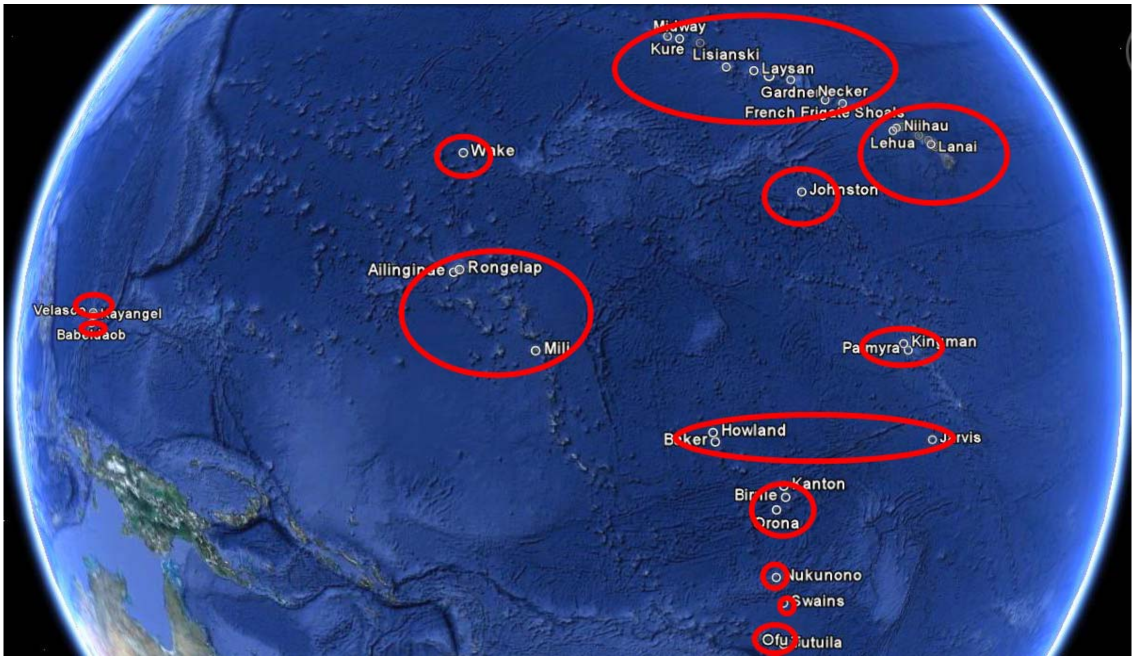

Figure 2.6: Map of the study area with the thirteen clusters identified at the similarity index level of 60 circled in red.

Detailed description of island groups:

1. Johnston Atoll: This group is made up of a single isolated atoll located to the Southwest of the Hawaiian Island chain. It is quite a unique location with a low species richness (120 total species recorded in this study). Johnston could be considered a transitional location from the major source of immigration to the Hawaiiain Island chain from islands further West and South as there are very few species found in islands west and south as well as in the Hawaiian Islands that are not found at Johnston. Johnston is also interesting in its high abundance of hybrid species. Several wrasses (*Thalassoma duperrey, T. lutescens* and *T. quinquevittatum*) as well as certain butterflyfish species (*Chaetodon auriga* and

C. trifascialis) were observed that seemed to show characteristics of some mixture of these closely related species.

2. Main Hawaiian Islands: Eight islands clustered into this group that is made up entirely of volcanic high islands in the Southeast portion of the Hawaii chain. There is a great variety of islands in this group in terms of sizes ranging from the large islands of Oahu and Maui down to very small islands like Molokini and Lehua (note that Hawaii island – the largest in this group - was not surveyed in this study). The faunal similarities of the Main Hawaiian Islands are no surprise as they share many endemic species (approximately 25% of all known Hawaiian coral reef fish are endemic species). The species richness of the Main Hawaiian Islands is considerably higher than that of the Northwest Hawaiian Islands. This could be due to a variety of factors including the availability of additional habitat types not found in the atolls of the Northwest Hawaiian Islands, temperature tolerances, or other factors.
3. Northwest Hawaiian Islands: This group is made up of the atolls of the Hawaiian Island chain along with the northernmost high island, Gardner Pinnacle. These atolls are the sunken remnants of once high islands that have eroded and subsided as the Pacific plate has moved Northwest over the course of the last 30 million years. All of these islands except Midway are uninhabited (with the exception of the occasional field biologist camping seasonally on several of the islands doing work on seals, turtles, or birds). The available habitat in atolls is quite limited in comparison to the high islands that make up the Main Hawaiian Islands to the south and east which likely accounts for the difference in species seen here. There are also a number of fish species known from deeper water in the Main Hawaiian Islands that were present within survey depths in the Northwest Hawaiian Islands presumably as they were able to live shallower in the water as the temperatures fall moving North. It is also worth noting that the Northwest Hawaiian Islands are the highest latitude reefs found in the Pacific and as such are likely subject to environmental conditions that are different from other coral reef areas.
4. Marshall Islands: The Marshall Islands are made up of 29 atolls and five islands and are situated north of the equator just over 2000 miles southwest of Honolulu.

Three atolls of the Marshalls were surveyed for this project: Ailinginae, Rongelap, and Mili. Ailinginae and Rongelap are in the northern end of the chain and Mili is at the southern end. Ailinginae has always been uninhabited and was used occasionally as a fishing ground by the people of Rongelap historically. Rongelap had a population of several hundred up until the island was impacted by radioactive fallout from the US nuclear testing program on the nearby atoll of Bikini. The island was evacuated, repopulated, and then evacuated again in 1985. It remained uninhabited until a small crew of construction workers returned to it just prior to this study. Mili is an atoll with a population of approximately 1000. The Marshall Islands are the Easternmost part of Micronesia and, along with Wake Atoll, represent the final stepping stone for species potentially reaching the Hawaiian Islands from the west.

5. Babeldaob: This is the largest island of the Republic of Palau. It is a high island with many different habitat types including a large lagoon, barrier reef, fringing reefs, channel areas, seagrass beds, and mangroves. It is an extremely diverse area in terms of fish species with 494 species recorded during this study. There are not many endemic species known from Palau – or any of the islands of Micronesia – as the islands of Micronesia are not very far apart when compared to the isolation of places like the Hawaiian Islands. The population of Babeldaob was approximately 6,000 at the time of this study. It is much less densley populated than the smaller island of Koror which lies to the south and which is the center for Palauan tourism and business.
6. Northern Palau: Kayangel Atoll and Velasco Reef are the northernmost islands of Palau. Kayangel is an uninhabited atoll while Velasco is a submerged seamount that rises to approximately 20 meters below sea level at its shallowest point. While very similar to Babeldaob, they did form their own cluster in the dendrogram largely due to the lack of nearshore species that inhabit the shorelines of Babeldaob.
7. American Samoa: One atoll – Rose Atoll- and four high islands – Ofu, Olosega, Tau, and Tutuila make up this cluster which is the furthest south of the study sites. Very small human populations exist on Ofu, Olosega, and Tau, while the vast

majority of the 65,000 residents of the territory live on Tutuila. Rose Atoll is uninhabited and lies approximately 150 miles east of Tutuila and was recently designated as a Marine National Monument by President George Bush.

8. Wake Atoll: Wake is a US military installation that lies 2,100 miles almost due west of Honolulu. It clusters alone and is quite different from the atolls of the Marshall Islands that lie due south. At the time of this study, there was a population of 90 military and civilian contractor personnel stationed on the island.
9. Nukunono Atoll: Part of Tokelau, Nukunono had a population at the time of this study of less than 500. It clusters separate from Swains Island which is described below and is the only other island of Tokelau included in this survey (Atafu and Fakaofo were not able to be visited due to time constraints).
10. Northern US Line Islands: This cluster consists of Palmyra Atoll and Kingman Reef. Kingman is an uninhabited emergent low island and Palmyra is an atoll with a small population of biologists (typically around twenty) residing on it. Palmyra has historically had a military presence on it and does include such human developments as a runway and has had significant human alteration to natural water flow due to construction that eliminated some of the channels that previously allowed lagoonal waters to circulate more fully.
11. US Phoenix Islands and Equatorial Line Islands: Howland and Baker Islands are very small islands that lie almost directly on the equator. Jarvis island is politically considered part of the US Line Islands, but based on this study, it clustered more closely with the US Phoenix Islands. Jarvis is also situated very near the equator, and while it is geographically closer to Palmyra and Kingman, the faunal similarities between these three equatorial islands suggest larval transport between them entrapped in the equatorial current.
12. Swain's Island: Like Nukunono (above), Swain's is geographically and was historically considered part of Tokelau. However, it is politically connected to American Samoa. It lies just over 200 miles to the North of Tutuila and about 140 miles Southeast of Nukunono. Geologically it is interesting as it is a low island with water in the middle, but it is not an atoll. The water in the center of the island is a brackish lake and there are no channels providing for surface flow to the

ocean. It therefore lacks a good portion of the habitat that Nukunono has in its extensive lagoon.

13. The Southern Phoenix Islands: Birnie, Kanton, and Orona are three atolls that make up the Southern Line Islands. Kanton had a population of a few dozen at the time of this study and the other two atolls were uninhabited.

A scatterplot of the Bray-Curtis similarity between island pairs versus the distance between them suggests a high degree of spatial autocorrelation (Fig. 2.7).

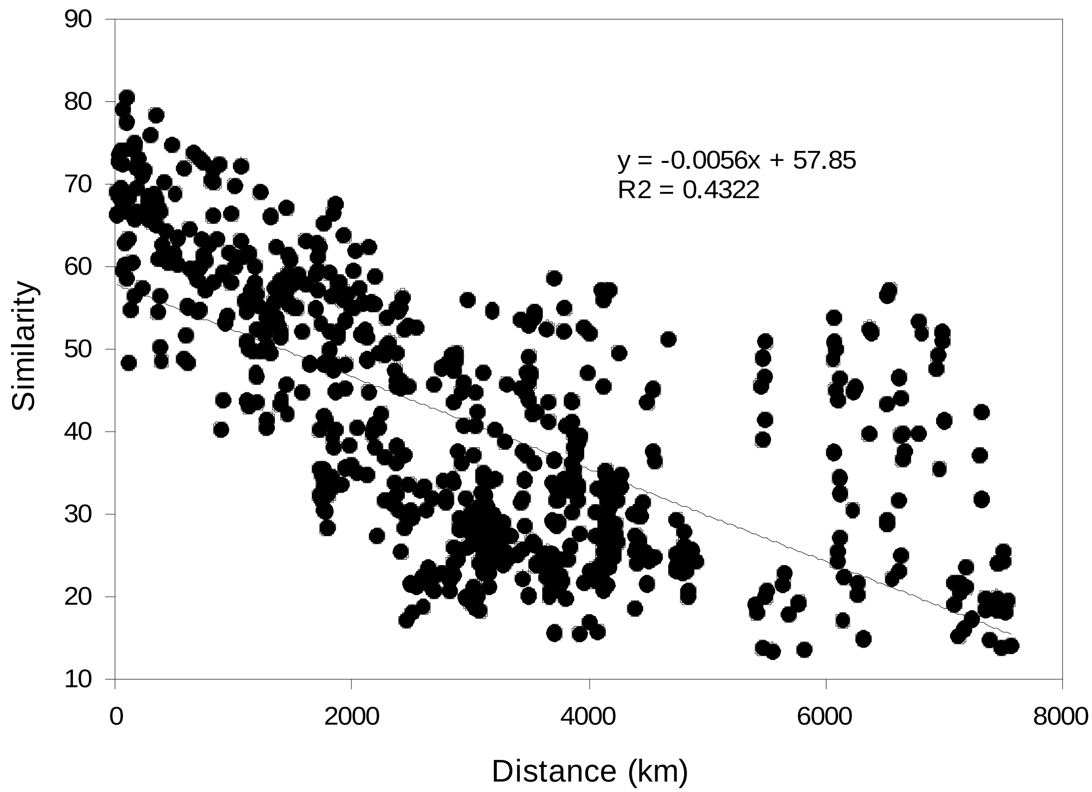

Figure 2.7: Scatterplot and regression line of Bray-Curtis similarity index versus geographic distance between island pairs.

To explicitly test this hypothesis, a Mantel test was run comparing geographic distance between island pairs to the level of their similarity (measured as Raup-Crick dissimilarity). The resulting simulated p-value of 0.0004 allows us to reject the null hypothesis that there is no relationship between distance and similarity, and the positive coefficient of correlation (0.33) indicates that more distant island pairs are in fact more dissimilar.

Discussion:

The reality of visually surveying coral reef fishes means that none of the islands surveyed in this study were fully censused. No matter how fully sampled a coral reef environment is, there are always going to be fish species that are not included on visual census. Many cryptic species that live in the sand or under ledges will simply not be observed without using destructive sampling techniques. Certainly the exact quantitative nature of the relationships between islands depicted in this study would change if complete census methods were used, but the methods used throughout the study were consistent and as thorough as possible given the time constraints posed by the extremely remote locations of many of the study sites. Perhaps there is a higher level of local and/or regional endemism among the more rare and cryptic species that went unnoticed in this study that would strengthen or alter the clustering of site groups, but it is unlikely that the processes affecting the distribution of cryptic species are different enough from the processes affecting the distribution of well sampled species that their inclusion would make significant differences in the relationships between islands seen here.

As expected, strong geographic autocorrellation was found. The strength of the geographic cohesion of the clusters supports the idea that sufficient sampling was done to adequately characterize regional groupings of islands. The importance of this finding is that it lends support to the idea that reef fish communities can be sufficiently sampled using underwater visual techniques like those used in this study. As discussed in subsequent chapters of this dissertation, great care was taken in this study to insure that all data used was reliable and robust enough to merit the types of analyses that were taken. In attempting to answer questions with broad geographic relevance and that include a large number of fish species and families, there was initial concern over the ability for adequately robust data to be obtainable using these methods. This chapter represents the first step at addressing some of these concerns, while at the same time defining the regions that will be used to test the major hypothesis at the end of this study concerning the relation of population density to dispersal likelihood.

Further, the geographic cohesion of the clusters revealed here suggests that for any given island, there exists a regional collection of potential species that can populate each island within that region. While one could speculate that ecological differences between islands are less likely to be a determining factor in species composition than the chance arrival of a species, the unfortunate confounding problem of most closely situated islands being of similar geologic origin and environmental condition makes it difficult to test this notion. The clustering of Babeldaob in Palau distinct from the other 2 Palau study sites is one example in this study that seems to refute this notion, as the environmental differences between Babeldaob and the other two Palauan islands are extreme yet their locations are close. Another visible break in the clustering of closely situated islands is apparent between the high islands of the Main Hawaiian Islands and the atolls of the Northwest Hawaiian Islands. Clearly different environmental conditions exist between the high islands and the low-lying atolls, and the clustering based on fish species composition reflects this split clearly. In most of the other regions, clustering seems to be very much reflective of geographic proximity. However, while closely situated islands may have very similar species composition in terms of presence/absence checklists, there may certainly be ecological differences between islands that determine how successful one species is relative to another at any given island and as such, quantitative differences in community structures may be predicted. This idea is given further attention in chapters 2 and 3 of this dissertation.

The major break in the dendrogram isolating the Hawaiian Islands from the rest of the study areas was expected as Hawaii is known to be home to a high number of endemic species and known to be missing many of the species (and even families) that are widespread throughout the Pacific. The large distance that separates the Hawaiian Islands from the rest of the Pacific presents one of the largest physical gaps between islands in the Pacific. The likelihood of any coral reef fish species successfully migrating between the Hawaiian Islands and other islands or continents is extremely low, and as such, there has likely been a high number of new species formed in the Hawaiian Islands.

However, there does seem to be some level of gene flow between Hawaii and some of the islands surveyed in this study as there are a number of shared species across all study regions. This raises the question, what allows certain species to be able to successfully cross this gap while others can not? The intermediate positioning of Johnston Atoll on the outskirts of the Hawaiian Islands clusters suggests it is an intermediate step between Hawaii and the rest of the Pacific and represents the last stepping stone for species to make the jump across this oceanic gap. The importance of Johnston as a source for marine species entering the Hawaiian archipelago is reported for corals (Maragos et al. 2004) and has been suggested for fish (Rivera et al. 2004). Additionally, oceanographic modeling work has identified several currents that push larvae towards the Hawaiian archipelago from Johnston and which serve as the conduits by which the Hawaiian Islands have likely been colonized (Kobayashi 2006). The importance of the pattern found in this investigation is that it establishes the framework for testing the main hypothesis of this dissertation as it relates to the likelihood of a fish species making the jump to Johnston Atoll – gateway to Hawaii. The identification of the four main clusters of islands that form a semi-circle around Johnston allows for the testing of this hypothesis against Johnston's four most closely situated and most closely related island clusters in the Pacific.

CHAPTER 3:
PRACTICALITIES AND LIMITATIONS IN ASSESSING REEF FISH POPULATIONS USING VISUAL SURVEY METHODS

Abstract:

An assessment was conducted on the validity of several different types of visual census methods in robustly sampling coral reef fish populations. In an effort to restrict further analyses to only the most robust types of fish population data, several steps were taken to rigorously test and then restrict the data types that could be reliably used. To look at the ability to robustly assess fish populations in space, a new method was developed to test the strength of each of four data types typically collected during visual fish surveys. Based on this test which examines the geographic cohesion of clusters of sites situated around an island, we determined our best data types included numerical density estimates based on belt-transects (BLT-ND) and presence/absence records based on a thorough roving diver survey (RDS-PRES). These two data types produced greater than expected geographic cohesion when compared to a randomly generated set of site relationships, yet the clusters of sites based on the two data types consisted of different groups of sites and would lead an investigator to develop different mechanistic explanations for the resulting clusters. Significantly greater than expected geographic cohesion was produced using biomass density estimates from belt-transects (BLT-BD), but this data type obscured potentially important factors affecting community structure. Presence/absence records limited to records of species observed within the confined boundaries of belt-transects (BLT-PRES) showed less cohesion than expected and is considered a poor tool to assess populations.

Introduction:

The development of multivariate techniques to investigate ecological communities has enabled scientists to examine complex relationships between study sites based on large matrices of environmental variables (James & McCulloch 1990; Legendre &

Legendre 1998). These tools have become valuable for natural resource managers as increased attention is being paid to ecosystem-based management (Link et al. 2002). Taking advantage of the often large and complex databases of ecological information that can be collected from a single study area, multivariate methods have greatly increased our ability to define communities within an ecosystem – a necessary prerequisite to developing conservation priorities or evaluating the effects of natural disturbances on reef fish communities (Booth & Beretta 2002). In addition to providing baseline reef fish monitoring data at a location that had previously been largely unstudied, this investigation addresses methodological questions regarding the surveying and subsequent multivariate analysis of reef fish communities.

From October 18 to 22, 2005 scientists from the NOAA Coral Reef Ecosystem Division conducted the first quantitative assessment of the shallow water coral reef fishes of Wake Atoll. Wake Atoll is located at 19˚ 18' N, 166˚ 36' E and is currently under the control of the United States Army. At the time of our visit, it was inhabited by a population of approximately 200 workers who are stationed to maintain the infrastructure and defense operations. The atoll consists of 3 main islets – Peale in the North, Wilkes in the Southwest, and Wake in the South and East (Fig. 3.1). Several scientific collections of fishes have been made at the island beginning in 1925 (Fowler & Ball, 1925). Most recently, SCUBA surveys were conducted by Lobel and Lobel (2004) who published an excellent review of previous scientific work at Wake as well as the most recent species list. Our work represented the first attempt to quantitatively assess the shallow water reef fish populations at this remote and largely inaccessible coral atoll.

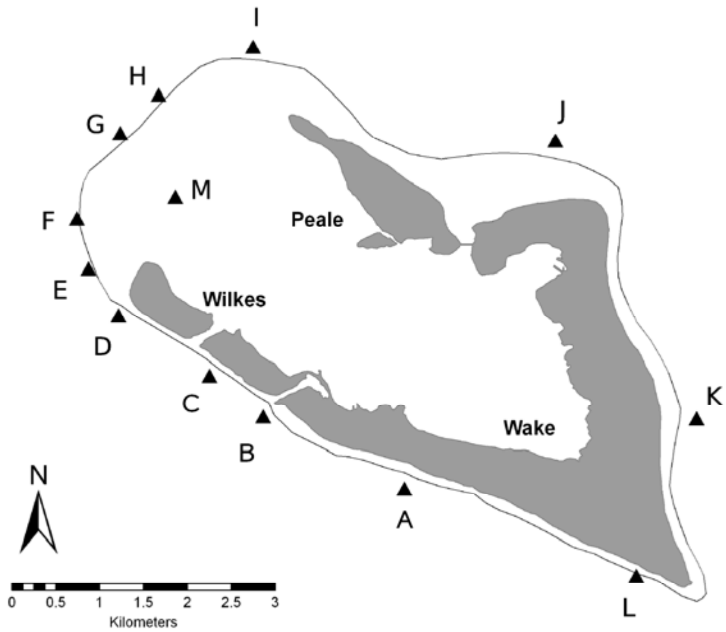

Fig. 3.1: Map of Wake Atoll showing the locations of survey sites. Shaded area indicates emergent land. Solid line follows the reef crest.

Lobel and Lobel (2004) performed a thorough qualitative census and recorded many previously undocumented species. They identified four basic aquatic habitat types at the atoll: the shallow, turbid lagoon; the intertidal reef flat; the reef crest; and the outer reef "which drops rapidly in depth on both the exposed windward and sheltered leeward sides." They made no further attempt to classify fish communities within these habitat types or to examine differences between parts of the atoll. Our efforts were focused almost exclusively in the fourth habitat type as the primary interest in this investigation is to examine a single habitat type over an extremely wide scale.

Here, we attempt to define the relationships that exist between reef slope sites around the atoll based on qualitative as well as several measures of quantitative fish species composition. Additionally, we evaluate the strength of each type of data collected in terms of how well site relationships based on each data type compared to randomly generated matrices of site relationships. In this endeavor, we created a new method and framework for evaluating the strength of each data type for effectively classifying distinct communities.

Methods:

A combination of several widely used methods in sampling reef fish were used to characterize the fish fauna at Wake Atoll. First, fish were surveyed using a modified belt-transect method (BLT) following Friedlander & DeMartini (2002). Two divers swam in tandem along a 25m transect line recording abundance and size of all fish >20cm TL observed within a 4m wide belt on their side of the transect line. The divers then returned along the line recording abundance and size of all fish <20cm TL observed within a 2m wide belt on their side of the transect line. All fish were identified to the lowest possible taxon which was the species level in nearly all cases. Sizes were estimated based on visual estimates and were to the nearest 1 cm for fish between 1 and 5 cm in total length and to the nearest 5 cm for fish larger than 5 cm. Three transects separated by 5m were performed at each site. For the analyses, data from the three transects were summed and treated as a single site. Following completion of the transects, the two divers performed a roving diver survey (RDS) recording the presence of all species observed at the site. BLT surveys were conducted at a standard depth (~15 m) at all sites. RDS surveys were restricted to the safe diving limits imposed by the NOAA dive safety office and typically ranged from 35 m up to the surge zone.

Sites were selected haphazardly to give the best possible spatial coverage around the atoll attempting to survey at regularly spaced intervals around the entire circumference of the atoll. At each site, a predetermined GPS coordinate was located where divers entered the water and then swam down the reef slope to the prescribed 15m depth where the transect began. Transects were run along a constant depth contour and began in an upcurrent direction. Completion of the three transects at each site took approximately 45 minutes. The RDS surveys lasted for an additional 15-20 minutes at each site. Weather conditions and diver safety concerns precluded surveys along portions of the exposed north and east coasts. Twelve sites on the outside of the atoll and one site within the lagoon were surveyed. The lagoon is a severely disturbed habitat and the fish fauna observed there differed greatly from the fish observed on the outside. Species presence and densities from the one lagoon site were recorded, but this site was not included in any of the cluster analyses or multivariate investigations.

The surveys yielded four types of data at each site. From the belt transects, numerical density (BLT-ND) of each species was calculated and a list of species present within the boundaries of the belt transects was recorded (BLT-PRES). Additionally, the size class information was used to calculate biomass estimates for all fish observed, and the biomass density (BLT-BD) was calculated for each species at each site. Biomass calculations were based on published length-weight relationships for each species (www.Fishbase.org). When no information for a species was available, a proxy was used from a closely related species or a mean for the family/genus. Finally, the roving diver surveys yielded a more complete list of species present at each site (RDS-PRES). The RDS-PRES data is essentially the same type of data collected by Lobel and Lobel (2004), except that they additionally performed a few chemical collections of small, cryptic fishes that are not typically observed by divers.

To investigate relationships between sites, multivariate cluster analysis was performed using PRIMER-E®, version 6 (Clarke & Warwick 2005). Clusters were based on Bray-Curtis similarity matrices and were created using the complete linkage cluster mode. BLT-ND and BLT-BD data were fourth-root transformed before creating the similarity matrices to reduce the weight of the most abundant species and to increase the weight of rare species in the analyses (Clarke & Warwick 2005).

To further test the sensitivity of the data types to the technical aspects of the clustering procedures, analyses were repeated using the group average cluster mode as well.

Following the cluster analyses, SIMPER analyses were also performed on 4^{th} root transformed BLT-ND and BLT-BD data at all cluster levels to determine which fish species contributed most heavily to the similarities within the clusters.

To examine the strength of each data type, a new method was devised to analyze the geographic cohesion of clusters resulting from analysis of each data type. A grouping in a cluster analysis (i.e. a limb in a dendrogram) was defined as "geographically cohesive" if the sites included in the group were located contiguously along the outside of the atoll rim. Clusters consisting of a single site were considered geographically cohesive. Note that geographic cohesion in this sense does not imply that nearby sites necessarily have similar communities, since two spatially adjacent cohesive clusters could be quite dissimilar (i.e. far apart in the cluster diagram). Our analysis therefore differed from the common assessment of correlation between geographic and ecological distance. Further, this analysis required special consideration of relative site positions as the sites were situated on a ring and were not laid out in a linear manner as would be experienced if the investigation were performed at a collection of study sites along a coastline of broader scale.

Each dendrogram produced by the cluster analyses was examined by counting the geographically cohesive clusters at each level of the tree (i.e. below each node). These counts were then compared to the distributions of counts expected under the null hypothesis of no geographical cohesiveness. Two kinds of null models were used: one created completely random trees for the 12 geographic locations, and the second constrained the topology of the tree to that actually observed for that data type. To generate a random tree, random numbers were assigned to the 12 sites and cluster analysis was performed on these "data." A total of 100,000 such trees were generated and for each the number of cohesive clusters was counted (at each level of the tree and in total) and tabulated. This distribution of frequencies of cohesive clusters was used for comparison to the counts observed for all four data types. The "constrained" null model was simulated by randomizing the assignment of site identities to branch tips of the tree, and counting cohesive clusters. A total of 10,000 such randomizations were performed for each of the four trees.

These analyses of geographic cohesion were performed using R (R Development Core Team, 2004). Analyses of random trees used the clustering function "agnes" in the R package "cluster" and the function "cutree" in the "stats" package.

Results:

A total of 184 fish species from 36 families were recorded. Of these, 22 species represented new records from Wake Atoll, and one has been identified as a new species (Randall et al., in prep). The fish fauna was dense as expected in an area with such little fishing pressure. Table 3.1 lists all species recorded during this survey including our new records. Numerically, the most abundant species were three pomacentrids, *Chromis acares, C. agilis,* and *C. vanderbilti* that together accounted for 57% percent of all fish observed. In terms of biomass, a parrotfish, *Chlorurus frontalis*, the grey reef shark, *Carcharhinus amblyrhynchos*, and the Napoleon wrasse, *Cheilinus undulatus*, were the three species with the greatest biomass, accounting for 53% of the total fish biomass.

Several subjective impressions stand out from these surveys. First, there appeared to be little change in species composition with depth across our diving depth range. Few deep water species were observed along the steep outer reef dropoffs, and at each site, the species composition appeared relatively homogeneous between 3 and 35 m depths which typically corresponded to the depth limits of the roving diver surveys. Second, typical fishery target species were unusually large – especially in girth. This seemed particularly true among parrotfish. We speculate that many of the biomass estimates used in this investigation are underestimates, as our conversion from length to weight were based on published reports of relationships calculated from samples collected in other locations. Unfortunately, collections of these species from Wake have not yet been conducted. Finally, many fish that are severely depleted throughout their ranges were exceptionally abundant at Wake.

Table 3.1: List of all fish species recorded at Wake Atoll. Asterisks indicate new records at this location.

Family	Species
Carcharhinidae	
	Carcharhinus amblyrhynchos
	Galeocerdo cuvier *
Myliobatidae	
	Aetobatus narinari
Muraenidae	
	Gymnothorax javanicus
Holocentridae	
	Myripristis adusta
	Myripristis berndti
	Neoniphon sammara
	Sargocentron spiniferum
	Sargocentron tiere
Syngnathidae	
	Corythoichthys flavofasciatus
Aulostomidae	
	Aulostomus chinensis
Fistulariidae	
	Fistularia commersonii
Scorpaenidae	
	Pterois antennata
Serranidae	
	Cephalopholis argus
	Cephalopholis spiloparaea
	Cephalopholis urodeta
	Epinephelus fasciatus
	Epinephelus hexagonatus
	Epinephelus howlandi *
	Epinephelus maculatus *
	Epinephelus melanostigma *
	Epinephelus merra
	Epinephelus polyphekadion
	Epinephelus sp (Serranidae)
	Epinephelus tauvina
	Pseudanthias pascalus
Cirrhitidae	
	Neocirrhites armatus
	Paracirrhites arcatus
	Paracirrhites forsteri
	Paracirrhites hemistictus
Apogonidae	
	Cheilodipterus quinquelineatus
Malacanthidae	
	Malacanthus brevirostris

Carangidae

Carangoides ferdau
Carangoides orthogrammus
Caranx lugubris
Caranx melampygus
Caranx sexfasciatus
Scomberoides lysan
Selar crumenophthalmus *
Seriola dumerili *
Trachinotus baillonii

Lutjanidae

Aphareus furca
Aprion virescens
Lutjanus fulvus
Lutjanus monostigma

Lethrinidae

Lethrinus obsoletus
Lethrinus olivaceus *
Lethrinus rubrioperculatus
Monotaxis grandoculis

Mullidae

Mulloidichthys flavolineatus
Mulloidichthys vanicolensis
Parupeneus barberinus
Parupeneus cyclostomus
Parupeneus insularis
Parupeneus multifasciatus

Pempheridae

Pempheris oualensis

Chaetodontidae

Chaetodon auriga
Chaetodon ephippium
Chaetodon lunula
Chaetodon ornatissimus
Chaetodon punctatofasciatus
Chaetodon quadrimaculatus
Chaetodon reticulatus
Chaetodon semeion
Chaetodon trifascialis *
Chaetodon ulietensis
Chaetodon unimaculatus
Forcipiger flavissimus
Forcipiger longirostris
Hemitaurichthys thompsoni
Heniochus acuminatus

Pomacanthidae

Centropyge flavissima
Centropyge loricula
Centropyge multifasciata *

Kyphosidae

Kyphosus bigibbus
Kyphosus cinerascens
Kyphosus sp (Kyphosidae)
Kyphosus vaigiensis *

Oplegnathidae

Oplegnathus punctatus

Pomacentridae

Abudefduf sordidus
Abudefduf vaigiensis
Chromis acares
Chromis agilis
Chromis vanderbilti
Chrysiptera glauca
Dascyllus aruanus
Plectroglyphidodon dickii
Plectroglyphidodon imparipennis
Plectroglyphidodon johnstonianus
Plectroglyphidodon lacrymatus
Plectrogyphidodon phoenixensis
Stegastes fasciolatus
Stegastes lividus (probably misidentification)
Stegastes sp (Pomacentridae)

Labridae

Anampses caeruleopunctatus
Anampses melanurus *
Bodianus anthoides
Cheilinus chlorourus
Cheilinus trilobatus
Cheilinus undulatus
Coris aygula
Epibulus insidiator
Gomphosus varius
Halichoeres melasmapomus *
Halichoeres ornatissimus
Halichoeres trimaculatus
Hemigymnus fasciatus
Labroides bicolor
Labroides dimidiatus
Labropsis xanthonota *
Novaculichthys taeniourus
Oxycheilinus bimaculatus *
Oxycheilinus unifasciatus
Pseudocheilinus hexataenia
Pseudocheilinus octotaenia
Pseudocheilinus tetrataenia
*Pseudojuloides sp. *(New Species)*
Stethojulis bandanensis
Thalassoma lunare *
Thalassoma lutescens
Thalassoma purpureum

Thalassoma quinquevittatum
Thalassoma trilobatum

Scaridae

Bolbometopon muricatum
Calotomus carolinus *
Chlorurus frontalis
Chlorurus microrhinos
Chlorurus sordidus
Hipposcarus longiceps
Scarus altipinnis *
Scarus forsteni
Scarus frenatus *
Scarus ghobban
Scarus globiceps
Scarus oviceps
Scarus psittacus
Scarus rubroviolaceus
Scarus sp

Blenniidae

Cirripectes variolosus

Gobiidae

Amblygobius phalaena
Coryphopterus sp.
Gobiidae sp.

Siganidae

Siganus argenteus

Zanclidae

Zanclus cornutus

Acanthuridae

Acanthurus achilles
Acanthurus blochii
Acanthurus guttatus
Acanthurus leucopareius
Acanthurus nigricans
Acanthurus nigricauda
Acanthurus nigrofuscus
Acanthurus nigroris
Acanthurus nubilus
Acanthurus sp.
Acanthurus thompsoni
Acanthurus triostegus
Ctenochaetus cyanocheilus
Ctenochaetus hawaiiensis
Naso annulatus *
Naso brevirostris
Naso hexacanthus
Naso lituratus
Naso unicornis
Naso vlamingii
Zebrasoma flavescens

Zebrasoma veliferum

Sphyraenidae

Sphyraena barracuda

Bothidae

Bothus mancus

Balistidae

Balistoides viridescens
Melichthys niger
Melichthys vidua
Odonus niger *
Rhinecanthus rectangulus
Sufflamen bursa
Xanthichthys mento

Monacanthidae

Aluterus scriptus
Cantherhines dumerilii

Ostraciidae

Ostracion cubicus
Ostracion meleagris

Tetraodontidae

Arothron meleagris
Arothron nigropunctatus *
Canthigaster amboinensis
Canthigaster solandri

Diodontidae

Diodon hystrix

Species of Concern

Due to their severely depleted and potentially overexploited status, two Pacific coral reef fish are recent additions to the National Marine Fisheries Service list of Species of Concern (NOAA 2006). Both of these species, the Humphead or Napoleon Wrasse, *Cheilinus undulatus*, and the Bumphead Parrotfish, *Bolbometopon muricatum*, were notably abundant at Wake. Napoleon Wrasse were observed at 9 of 13 sites surveyed. Groups of up to seven individuals were observed and the largest individuals were estimated to be 120 cm TL. Bumphead Parrotfish were observed at 5 of 13 sites surveyed. Large herds of these important grazers were often seen cruising along the reef crest, but they would typically keep their distance from divers and our transect areas. On the outside reef, the largest schools observed were approximately 30 individuals, and they ranged in size up to 75 cm TL. Inside the lagoon, they were not recorded at our one survey site. However, during an exploratory transit through the Eastern side of the lagoon, we observed upwards of 200 individuals in water as shallow as 2 feet over a sand substrate. Such a large concentration of this species is a unique situation that surely deserves further attention from conservationists and scientists.

Neither Napoleon wrasse nor Bumphead Parrotfish were recorded in great numbers along transects. It is uncommon for these species to approach within such narrow boundaries and approach divers. Yet, the frequency of their observation was exceptionally high and exceeds my experience at any other location surveyed in the Pacific Ocean.

Cluster Analysis Results

Based on the four data types, the dendrograms show strikingly different tree topologies and site relationships (Fig. 3.2). The BLT-ND analysis initially separates sites along the southern, eastern, and northeastern exposures of the atoll (sites L, J, K, A, B, C,

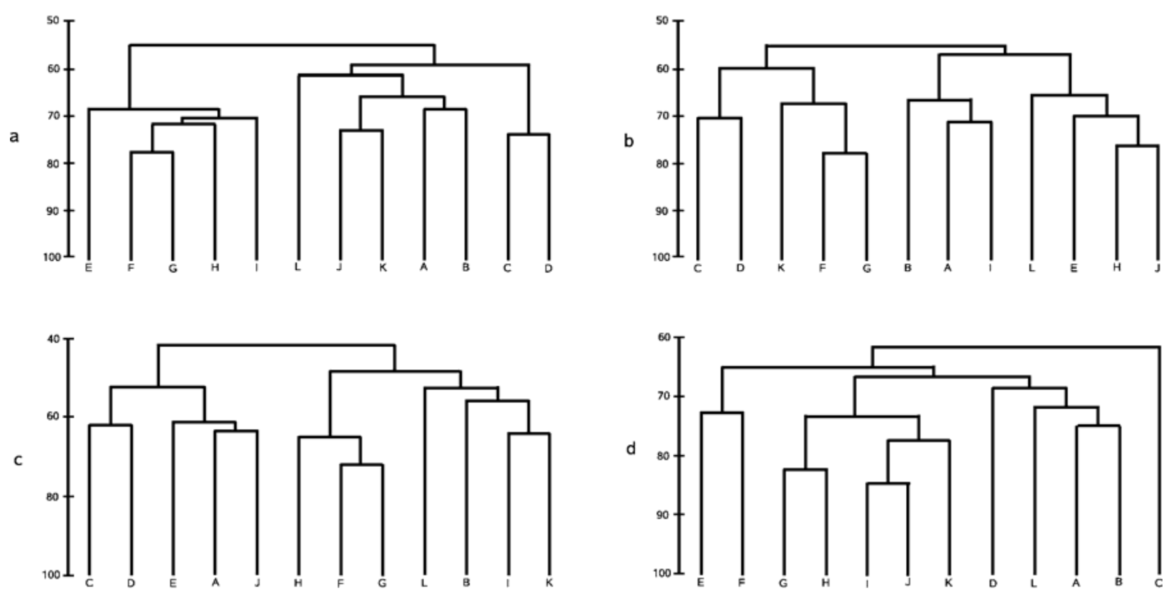

Fig. 3.2: Results from cluster analysis based on four data types: a. BLT-ND; b. BLT-PRES; c. BLT-BD; and d. RDS-PRES. Vertical axis indicates similarity.

and D) from sites along the northwest exposure (sites E, F, G, H and I). The group of sites from the south, east, and northeast is further broken down into clusters made up of sites from the western half of the southern exposure (sites C and D), sites from the eastern half of the southern exposure (sites A and B), sites from the northeast exposure (sites J and K) and the single site near the southeastern tip (site L). The cluster initially composed of the northwestern exposure sites sheds individual sites as one progresses along the branches of the dendrogram towards levels of increasing similarity. Sites are shed from alternating sides of the outer edges of this cluster until the two most similar sites at the atoll (sites F and G) are the last remaining multi-site cluster. This data type shows high geographic cohesion at every level of clustering. The only non-cohesive cluster arises at the 4-cluster level when site L (the uniquely situated southeastern tip site) clusters separately from the cluster containing the two sites immediately adjacent in a clockwise direction around the atoll (sites A and B) and the two sites immediately adjacent in a counterclockwise direction (sites J and K).

The dendrogram from the BLT-PRES data does not show nearly as much geographic cohesion as the BLT-ND analysis. Clusters based on this analysis more typically include several non-contiguous locations around the atoll rim. Even among the three most closely related pairs of sites (sites F and G, sites H and J, and sites A and I), only one pair consists of adjacent sites. Summed over all levels of clustering, the BLT-PRES data shows only 44 cohesive clusters, the lowest total of any of the data types used in this investigation.

The BLT-BD analysis reveals an intermediate level of cluster cohesion with 50 of the 65 possible clusters being geographically cohesive. There is a lack of cohesion at the base of the tree (neither of the first two clusters is cohesive) and an increasing frequency of cohesion as one moves towards increasing numbers of smaller clusters.

The final data type, RDS-PRES, shows levels of cohesion similar to those of the BLT-ND data (63 out of 65 clusters are geographically cohesive). However, the dendrogram shows a strikingly different topology, and the resultant clusters are different than the BLT-ND clusters. Proceeding from least similarity to increased similarity, the first split in the tree separates one site (C) from all other sites at the atoll. The next split extracts the two contiguous sites (E and F) from the main cluster leaving a large, non-cohesive cluster containing sites G, H, I, J, K, D, L, A, and B.

An example of geographic cohesion at a single cluster level across all four data types is shown in Figure 3.3. Taken from the 5-cluster level, this figure shows cluster cohesion in a diagrammatic map of the atoll. The BLT-ND data and the RDS-PRES data show complete cohesion of all clusters and all sites at this level, while the BLT-BD and BLT-PRES data show dramatically less cohesion.

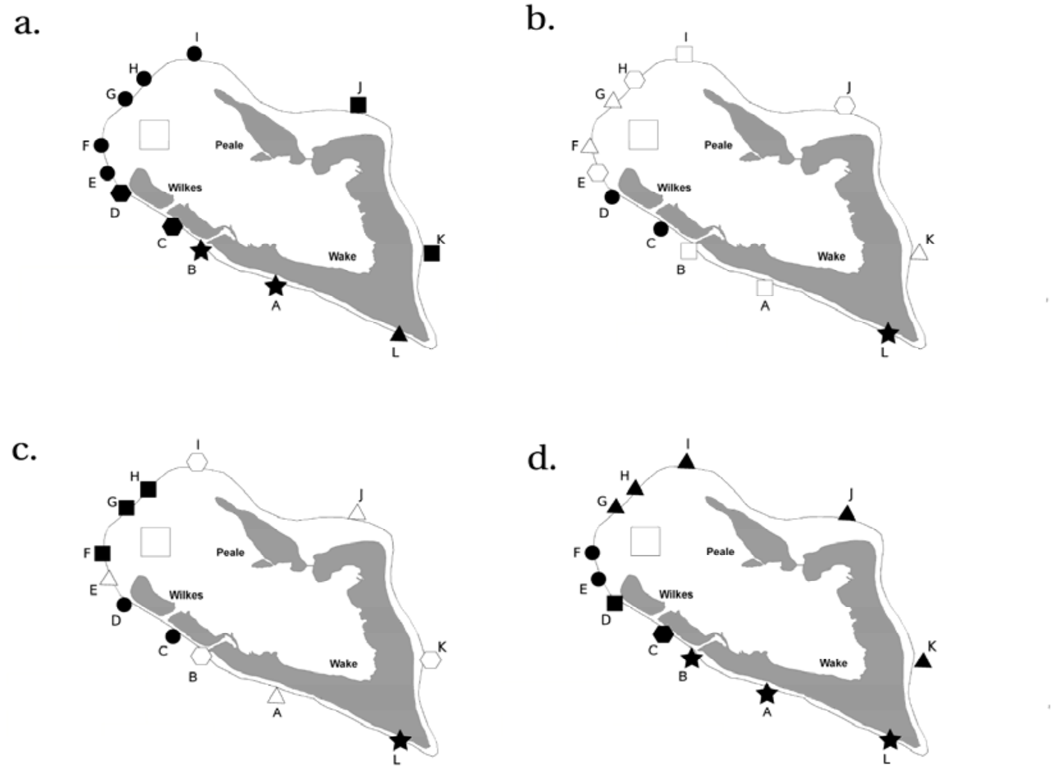

Fig. 3.3: Spatial representations of cluster cohesion from the four data types at the 5-cluster level. Sites which cluster together at this level of clustering share one of the five symbols: circle, triangle, square, star, or hexagon. Symbols for clusters which are cohesive are filled in black. Symbols for non-cohesive clusters are not filled. The four data types are shown in the same arrangement as in Figure 2: a. BLT-ND; b. BLT-PRES; c. BLT-BD; and d. RDS-PRES.

The RDS-PRES data includes 65 species not observed within the boundaries of belt transects (Tab. 3.2). A mean of 81 species were recorded at each site with 48 of them being observed within the BLT boundaries. These numbers increase slightly to 86 and 50 if the lagoon site M is not included.

Site	RDS	BLT
A	79	50
B	81	51
C	75	39
D	79	49
E	77	51
F	74	55
G	90	48
H	89	55
I	93	51
J	84	50
K	98	46
L	96	55
M	40	19
Mean	81.2	47.6
w/o site M		
Mean	84.6	50.0

Tab. 3.2: Number of species recorded at each site within the boundaries of the belt transects (BLT) and beyond these boundaries during the roving diver survey (RDS).

The simulation results based on the test of cohesive clusters expected from a random tree of twelve sites is shown in Table 3.3. Summed over all levels of clustering, the maximum number of cohesive clusters possible is 65. The threshold for significance at the p = 0.05 level was 48 total cohesive clusters. This threshold was met by three of the data types. The BLT-ND, RDS-PRES and BLT-BD data clusters produced totals of 64, 63, and 50 cohesive clusters respectively. The BLT-PRES data produced only 44 cohesive clusters – less than expected at random. The number of cohesive clusters produced by the four data types was significantly greater than expected under the null hypothesis for the BLT-ND (for total numbers of cohesive clusters, one-sided tests, $p < 0.001$), RDS-PRES ($p < 0.001$), and BLT-BD ($p < 0.05$) data types; it was not statistically significant for the BLT-PRES data ($p > 0.05$) (Tab. 3.4a). Clusters were significantly more cohesive than random at all levels of the trees (except the very base and tips, where statistical tests are not feasible) for the BLT-ND and RDS-PRES data. Although the total number of cohesive clusters was significantly greater than expected for the BLT-BD data, the counts at different levels of the tree were not significant except at the five cluster level.

Number of Clusters	Number of Cohesive Clusters											
	0	1	2	3	4	5	6	7	8	9	10	11
2	0.945		0.055									
3	0.708	0.277	0.013	0.003								
4	0.353	0.502	0.135	0.009	0.001							
5	0.082	0.396	0.405	0.109	0.009	0.0004						
6	0	0.093	0.370	0.394	0.126	0.013	0.001					
7	0	0	0.052	0.310	0.132	0.182	0.023	0.001				
8	0	0	0	0	0.182	0.468	0.294	0.053	0.003			
9	0	0	0	0	0	0	0.378	0.476	0.136	0.009		
10	0	0	0	0	0	0	0	0	0.603	0.360	0.037	
11	0	0	0	0	0	0	0	0	0	0	0.820	0.180

Tab. 3.3: Cluster cohesion predicted from simulation results of random trees of 12 sites. Entries are the estimated probabilities of each number of cohesive clusters at each cluster level. Results are from a run of 100,000 random trees and are rounded to three digits.

Noting dramatic differences between the topologies of the four dendrograms, we also conducted more conservative randomization tests which assumed the particular topology observed for a given data type. The results were consistent with those obtained for the unconstrained trees (Tab. 3.4b). The observed degree of geographic cohesion was much greater than expected, in total and at all levels of the tree, for the BLT-ND and RDS-PRES data (for total numbers of cohesive clusters, one-sided tests, $p < 0.001$). The observed total number of cohesive clusters also was significantly greater than expected for the BLT-BD data ($p < 0.05$), but counts at different levels of the tree were not. The degree of geographic cohesion in the BLT-PRES tree again was not significantly different than expected at random ($p > 0.05$).

		Number of Clusters / Level of Tree									Total
		2	3	4	5	6	7	8	9	10	11	65
a.	BLT-ND	2	**3****	**3****	**5*****	**6*****	**7****	**8****	**9****	**10***	11	**64*****
	RDS-PRES	2	**2***	**3****	**5*****	**6*****	**7****	**8****	**9****	**10***	11	**63*****
	BLT-BD	0	1	2	3	4	5	6	8	**10***	11	**50***
	BLT-PRES	0	0	1	2	3	5	6	7	9	11	44
b.	BLT-ND	**2***	**3****	**3***	**5*****	**6*****	**7****	**8****	**9****	**10***	11	**64*****
	RDS-PRES	2	2	**3***	**5*****	**6****	**7****	**8****	**9***	**10***	11	**63*****
	BLT-BD	0	1	2	**3***	4	5	6	8	10	11	**50***
	BLT-PRES	0	0	1	2	3	5	6	7	9	11	44

*Tab. 3.4. The number of cohesive clusters for each data type at each level of clustering including statistical significance by comparison to (a) randomized trees with unconstrained topology and (b) randomized trees constrained to the same topology as the observed tree. * P < 0.05; ** P < 0.01; *** P < 0.001.*

Results from the SIMPER analyses of BLT-ND and BLT-BD data are shown in Tables 3.5a and 3.5b. A key to the four-letter species acronyms used in Tables 5a and 5b is included in Table 3.6. For all clusters at each cluster level, the three species that contributed most greatly to the similarity of sites within that cluster are shown. In most cases, these species were the three most abundant species in that cluster, however some exceptions occurred. Shown in the table is the mean numeric density (BLT-ND) or biomass density (BLT-BD), 4^{th} root transformed, of the top three species at all sites in the given cluster. For single-site clusters, the densities (4^{th} root transformed) of the three most abundant species are shown.

Tab. 3.5. Results from SIMPER analyses of (a) BLT-ND and (b) BLT-BD data. Shown in each cell are the sites included in that cluster, the three species contributing most heavily to the clusters' similarity, and the mean of the 4th root transformed density of those three species at all sites in the cluster. Organization of the sites follows the results from the cluster analyses (Fig. 3.2). Shading indicates a non-cohesive cluster. A key to the species codes is included as Tab. 3.6.

a.

<table>
<tr><th>Cluster Level</th><th colspan="12"></th></tr>
<tr><td>1</td><td colspan="12">C, D, E, A, J, H, F, G, L, B, I, K
CHFN 14.4
CAAB 12.6
CHUD 10.3</td></tr>
<tr><td>2</td><td colspan="5">C, D, E, A, J
SCFO 8.5
NALI 7.4
CHSO 6.9</td><td colspan="7">H, F, G, L, B, I, K
SCFO 10.1
NALI 9.4
CEAR 7.7</td></tr>
<tr><td>3</td><td colspan="5">C, D, E, A, J
SCFO 8.5
NALI 7.4
CHSO 6.9</td><td colspan="3">H, F, G
CHFN 20.2
NALI 11.1
SCFO 10.3</td><td colspan="4">L, B, I, K
SCFO 9.9
COAY 10.0
CEAR 7.4</td></tr>
<tr><td>4</td><td colspan="2">C, D
SCFO 9.1
NALI 7.3
MOGR 7.2</td><td colspan="3">E, A, J
CEAR 7.9
SCFO 8.1
NALI 7.4</td><td colspan="3">H, F, G
CHFN 20.2
NALI 11.1
SCFO 10.3</td><td colspan="4">L, B, I, K
SCFO 9.9
COAY 10.0
CEAR 7.4</td></tr>
<tr><td>5</td><td colspan="2">C, D
SCFO 9.1
NALI 7.3
MOGR 7.2</td><td colspan="3">E, A, J
CEAR 7.9
SCFO 8.1
NALI 7.4</td><td colspan="3">H, F, G
CHFN 20.2
NALI 11.1
SCFO 10.3</td><td>L
CHUD 18.2
SCFO 11.1
COAY 10.7</td><td colspan="3">B, I, K
SCFO 9.5
COAY 9.8
CEAR 7.6</td></tr>
<tr><td>6</td><td colspan="2">C, D
SCFO 9.1
NALI 7.3
MOGR 7.2</td><td colspan="3">E, A, J
CEAR 7.9
SCFO 8.1
NALI 7.4</td><td colspan="3">H, F, G
CHFN 20.2
NALI 11.1
SCFO 10.3</td><td>L
CHUD 18.2
SCFO 11.1
COAY 10.7</td><td>B
SECR 15.9
COAY 10.9
SCOV 9.3</td><td colspan="2">I, K
CAAB 13.8
SCFO 9.8
NALI 9.9</td></tr>
<tr><td>7</td><td colspan="2">C, D
SCFO 9.1
NALI 7.3
MOGR 7.2</td><td>E
CHFN 11.4
SCFO 9.3
CEAR 8.9</td><td colspan="2">A, J
COAY 8.2
SCFO 7.4
CEAR 7.5</td><td colspan="3">H, F, G
CHFN 20.2
NALI 11.1
SCFO 10.3</td><td>L
CHUD 18.2
SCFO 11.1
COAY 10.7</td><td>B
SECR 15.9
COAY 10.9
SCOV 9.3</td><td colspan="2">I, K
CAAB 13.8
SCFO 9.8
NALI 9.9</td></tr>
<tr><td>8</td><td>C
AENA 12.4
SCFO 9.0
NALI 7.5</td><td>D
CHMC 9.1
SCFO 9.1
CHSO 9.0</td><td>E
CHFN 11.4
SCFO 9.3
CEAR 8.9</td><td colspan="2">A, J
COAY 8.2
SCFO 7.4
CEAR 7.5</td><td colspan="3">H, F, G
CHFN 20.2
NALI 11.1
SCFO 10.3</td><td>L
CHUD 18.2
SCFO 11.1
COAY 10.7</td><td>B
SECR 15.9
COAY 10.9
SCOV 9.3</td><td colspan="2">I, K
CAAB 13.8
SCFO 9.8
NALI 9.9</td></tr>
<tr><td>9</td><td>C
AENA 12.4
SCFO 9.0
NALI 7.5</td><td>D
CHMC 9.1
SCFO 9.1
CHSO 9.0</td><td>E
CHFN 11.4
SCFO 9.3
CEAR 8.9</td><td>A
COAY 9.3
CEAR 8.0
SCFO 7.9</td><td>J
KYBI 9.0
MEVI 7.3
COAY 7.0</td><td colspan="3">H, F, G
CHFN 20.2
NALI 11.1
SCFO 10.3</td><td>L
CHUD 18.2
SCFO 11.1
COAY 10.7</td><td>B
SECR 15.9
COAY 10.9
SCOV 9.3</td><td colspan="2">I, K
CAAB 13.8
SCFO 9.8
NALI 9.9</td></tr>
<tr><td>10</td><td>C
AENA 12.4
SCFO 9.0
NALI 7.5</td><td>D
CHMC 9.1
SCFO 9.1
CHSO 9.0</td><td>E
CHFN 11.4
SCFO 9.3
CEAR 8.9</td><td>A
COAY 9.3
CEAR 8.0
SCFO 7.9</td><td>J
KYBI 9.0
MEVI 7.3
COAY 7.0</td><td colspan="3">H, F, G
CHFN 20.2
NALI 11.1
SCFO 10.3</td><td>L
CHUD 18.2
SCFO 11.1
COAY 10.7</td><td>B
SECR 15.9
COAY 10.9
SCOV 9.3</td><td>I
CAAB 14.2
SCFO 10.0
KYBI 10.0</td><td>K
CAAB 13.3
COAY 11.7
NALI 11.1</td></tr>
<tr><td>11</td><td>C
AENA 12.4
SCFO 9.0
NALI 7.5</td><td>D
CHMC 9.1
SCFO 9.1
CHSO 9.0</td><td>E
CHFN 11.4
SCFO 9.3
CEAR 8.9</td><td>A
COAY 9.3
CEAR 8.0
SCFO 7.9</td><td>J
KYBI 9.0
MEVI 7.3
COAY 7.0</td><td>H
CAAB 22.4
CHFN 19.4
KYBI 15.1</td><td colspan="2">F, G
CHFN 20.5
NALI 11.7
SCFO 10.2</td><td>L
CHUD 18.2
SCFO 11.1
COAY 10.7</td><td>B
SECR 15.9
COAY 10.9
SCOV 9.3</td><td>I
CAAB 14.2
SCFO 10.0
KYBI 10.0</td><td>K
CAAB 13.3
COAY 11.7
NALI 11.1</td></tr>
<tr><td>12</td><td>C
AENA 12.4
SCFO 9.0
NALI 7.5</td><td>D
CHMC 9.1
SCFO 9.1
CHSO 9.0</td><td>E
CHFN 11.4
SCFO 9.3
CEAR 8.9</td><td>A
COAY 9.3
CEAR 8.0
SCFO 7.9</td><td>J
KYBI 9.0
MEVI 7.3
COAY 7.0</td><td>H
CAAB 22.4
CHFN 19.4
KYBI 15.1</td><td>F
CHFN 17.6
NALI 11.9
CHUD 11.8</td><td>G
CHFN 23.4
KYBI 12.9
NALI 11.5</td><td>L
CHUD 18.2
SCFO 11.1
COAY 10.7</td><td>B
SECR 15.9
COAY 10.9
SCOV 9.3</td><td>I
CAAB 14.2
SCFO 10.0
KYBI 10.0</td><td>K
CAAB 13.3
COAY 11.7
NALI 11.1</td></tr>
</table>

Tab. 3.5b.

Cluster Level												
1	E, F, G, H, I, L, J, K, A, B, C, D CHAC 5.35 CHAG 3.62 CHVA 3.39											
2	E, F, G, H, I CHAC 5.82 CHAG 4.04 PSPA 3.70					L, J, K, A, B, C, D CHAC 4.78 CHVA 3.64 CHAG 3.07						
3	E, F, G, H, I CHAC 5.82 CHAG 4.04 PSPA 3.70					L, J, K, A, B CHAC 4.67 CHVA 3.71 CHAG 2.87					C, D CHAC 5.04 CHAG 3.58 CHVA 3.46	
4	E, F, G, H, I CHAC 5.82 CHAG 4.04 PSPA 3.70					L CHVA 4.93 CHAC 3.53 ACTR 2.99	J, K, A, B CHAC 4.95 CHAG 2.96 CHVA 3.40				C, D CHAC 5.04 CHAG 3.58 CHVA 3.46	
5	E, F, G, H, I CHAC 5.82 CHAG 4.04 PSPA 3.70					L CHVA 4.93 CHAC 3.53 ACTR 2.99	J, K CHAC 4.80 STFA 2.92 CHAG 3.02		A, B CHAC 5.10 CHAG 2.90 CHVA 3.16		C, D CHAC 5.04 CHAG 3.58 CHVA 3.46	
6	E CHAC 5.63 CHAG 4.24 PSPA 3.67	F, G, H, I CHAC 5.87 CHAG 3.99 PSPA 3.70				L CHVA 4.93 CHAC 3.53 ACTR 2.99	J, K CHAC 4.80 STFA 2.92 CHAG 3.02		A, B CHAC 5.10 CHAG 2.90 CHVA 3.16		C, D CHAC 5.04 CHAG 3.58 CHVA 3.46	
7	E CHAC 5.63 CHAG 4.24 PSPA 3.67	F, G, H, I CHAC 5.87 CHAG 3.99 PSPA 3.70				L CHVA 4.93 CHAC 3.53 ACTR 2.99	J, K CHAC 4.80 STFA 2.92 CHAG 3.02		A CHAC 5.65 CHAG 2.89 CHLU 2.69	B CHAC 4.56 CHVA 3.71 SECR 3.34	C, D CHAC 5.04 CHAG 3.58 CHVA 3.46	
8	E CHAC 5.63 CHAG 4.24 PSPA 3.67	F, G, H CHAC 5.90 CHAG 4.06 CHFN 3.44			I CHAC 5.77 CHAG 3.76 PSPA 3.74	L CHVA 4.93 CHAC 3.53 ACTR 2.99	J, K CHAC 4.80 STFA 2.92 CHAG 3.02		A CHAC 5.65 CHAG 2.89 CHLU 2.69	B CHAC 4.56 CHVA 3.71 SECR 3.34	C, D CHAC 5.04 CHAG 3.58 CHVA 3.46	
9	E CHAC 5.63 CHAG 4.24 PSPA 3.67	F, G CHAC 6.07 CHAG 4.16 PSPA 4.48		H CHAC 5.56 CHAG 3.87 MENI 3.76	I CHAC 5.77 CHAG 3.76 PSPA 3.74	L CHVA 4.93 CHAC 3.53 ACTR 2.99	J, K CHAC 4.80 STFA 2.92 CHAG 3.02		A CHAC 5.65 CHAG 2.89 CHLU 2.69	B CHAC 4.56 CHVA 3.71 SECR 3.34	C, D CHAC 5.04 CHAG 3.58 CHVA 3.46	
10	5 CHAC 5.63 CHAG 4.24 PSPA 3.67	F, G CHAC 6.07 CHAG 4.16 PSPA 4.48		H CHAC 5.56 CHAG 3.87 MENI 3.76	I CHAC 5.77 CHAG 3.76 PSPA 3.74	L CHVA 4.93 CHAC 3.53 ACTR 2.99	J CHAC 5.36 CHAG 3.15 STFA 2.93	K CHVA 4.42 CHAC 4.24 HAOR 3.30	A CHAC 5.65 CHAG 2.89 CHLU 2.69	B CHAC 4.56 CHVA 3.71 SECR 3.34	C, D CHAC 5.04 CHAG 3.58 CHVA 3.46	
11	E CHAC 5.63 CHAG 4.24 PSPA 3.67	F, G CHAC 6.07 CHAG 4.16 PSPA 4.48		H CHAC 5.56 CHAG 3.87 MENI 3.76	I CHAC 5.77 CHAG 3.76 PSPA 3.74	L CHVA 4.93 CHAC 3.53 ACTR 2.99	J CHAC 5.36 CHAG 3.15 STFA 2.93	K CHVA 4.42 CHAC 4.24 HAOR 3.30	A CHAC 5.65 CHAG 2.89 CHLU 2.69	B CHAC 4.56 CHVA 3.71 SECR 3.34	C CHAC 5.28 CHAG 3.71 CHVA 3.38	D CHAC 4.80 CHVA 3.53 CHAG 3.44
12	E CHAC 5.63 CHAG 4.24 PSPA 3.67	F CHAC 5.98 PSPA 5.38 CHAG 4.56	G CHAC 6.16 CHFN 4.06 CHAG 3.76	H CHAC 5.56 CHAG 3.87 MENI 3.76	I CHAC 5.77 CHAG 3.76 PSPA 3.74	L CHVA 4.93 CHAC 3.53 ACTR 2.99	J CHAC 5.36 CHAG 3.15 STFA 2.93	K CHVA 4.42 CHAC 4.24 HAOR 3.30	A CHAC 5.65 CHAG 2.89 CHLU 2.69	B CHAC 4.56 CHVA 3.71 SECR 3.34	C CHAC 5.28 CHAG 3.71 CHVA 3.38	D CHAC 4.80 CHVA 3.53 CHAG 3.44

	Acronym	Species	Family	Trophic Group	Max. Size	Social Structure
BLT-ND	ACTR	*Acanthurus triostegus*	Acanthuridae	Herbivore	17.5	Large aggregations
	CHAC	*Chromis acares*	Pomacentridae	Planktivore	7.5	Large aggregations
	CHAG	*Chromis agilis*	Pomacentridae	Planktivore	7.5	Large aggregations
	CHFN	*Chlorurus frontalis*	Scaridae	Herbivore	45	Small aggregations
	CHLU	*Chaetodon lunula*	Chaetodontidae	Benthic Invertebrates	17.5	Pairs/small aggregations
	CHVA	*Chromis vanderbilti*	Pomacentridae	Planktivore	5	Large aggregations
	HAOR	*Halichoeres ornatissimus*	Labridae	Benthic Invertebrates	12.5	Solitary/small aggregations
	MENI	*Melichthys niger*	Balistidae	Planktivore	25	Small/large aggregations
	PSPA	*Pseudanthias pascalus*	Serranidae	Planktivore	7.5	Large aggregations
	SECR	*Selar crumenophthalmus*	Carangidae	Benthic Invertebrates/Planktivore	30	Large aggregations
	STFA	*Stegastes fasciolatus*	Pomacentridae	Herbivore	7.5	Solitary
BLT-BD	AENA	*Aetobatus narinari*	Myliobatidae	Benthic Invertebrates	100	Small groups
	CAAB	*Carcharhinus amblyrhynchos*	Carcharhinidae	Piscivore/Large invertebrates	150	Solitary/small groups
	CEAR	*Cephalopholis argus*	Serranidae	Piscivore	40	Solitary
	CHFN	*Chlorurus frontalis*	Scaridae	Herbivore	45	Small aggregations
	CHMC	*Chlorurus microrhinos*	Scaridae	Herbivore	45	Small aggregations
	CHSO	*Chlorurus sordidus*	Scaridae	Herbivore	30	Small aggregations
	CHUD	*Cheilinus undulatus*	Labridae	Large invertebrates	120	Solitary/small groups
	COAY	*Coris aygula*	Labridae	Benthic Invertebrates	40	Solitary
	KYBI	*Kyphosus bigibbus*	Kyphosidae	Herbivore	30	Small/large aggregations
	MOGR	*Monotaxis grandoculis*	Lethrinidae	Benthic Invertebrates	40	Solitary/small aggregations
	NALI	*Naso lituratus*	Acanthuridae	Herbivore	40	Small/large aggregations
	SCFO	*Scarus forsteni*	Scaridae	Herbivore	40	Solitary/small aggregations
	SECR	*Selar crumenophthalmus*	Carangidae	Benthic Invertebrates/Planktivore	30	Large aggregations

Tab. 3.6. Key to the species acronyms used in Tab. 3.5. Max. size is the largest individual recorded on a transect during this survey.

Discussion:

A fundamental assumption of this investigation is that geographic cohesion is in fact present in ecological communities. The test of the various data types performed here relies entirely on this premise. While some critics may argue the merits of this assumption, it is the author’s belief, after performing thousands of fish transect surveys, that this assumption is well founded. Perhaps the strongest argument that this geographic cohesion is a real property of ecological communities lies in the fact that two of the data types tested here do in fact show this pattern of relatedness between sites. It is highly unlikely that such geographic cohesion as was revealed at every level of the cluster dendrograms based on the RDS-PRES and BLT-ND would be observed in two of the four data types if this was not a real property of the communities being studied. While perhaps circular in nature, this argument is worth considering by those who deny the premise of geographic cohesion to be a reasonable assumption.

Community classification is a fundamental foundation of natural resource management. Solid conservation plans require sufficient protection of as many representative habitats as possible (Ward et al. 1999). Often, a single data type is used in classifying these communities (Malcolm et al. 2012). This study vividly demonstrates that highly variable results can arise from analysis of different types of data collected from the same communities. Our attempt to test the strength of each data type as a tool to classify distinct ecological zones within a single atoll system serves as a caution, and hopefully a model to evaluate the strength of clustering methodologies.

Three data types showed significantly greater than expected geographic cohesion, yet the clusters resulting from the different data types were not concordant. To determine whether quantitative survey methods can provide additional insight into community structure when compared to presence-absence data, let us first examine the results from the RDS-PRES data which is essentially the same type of data used in the previous investigations of Wake. This data type showed very high levels of cluster cohesion. Proceeding from the base of the RDS-PRES cluster dendrogram, the least similar clusters are a single site cluster (site C) and a second cluster containing all the other sites. There are four fish species, *Zebrasoma flavescens, Coris aygula, Kyphosus bigibbus,* and *Parupeneus insularis*, that were not recorded at site C but were recorded at all other sites and there are three species, *Seriola dumerili, Anampses melanurus,* and *Odonus niger*, which were present at site C but absent from all other sites. Based on our observations, we can suggest no robust mechanism or set of environmental factors that would explain these findings. Two of the three species seen only at site C (*S. dumerili* and *A. melanurus*) are infrequently encountered by divers throughout their range, and the third species, (*O. niger*) is a common and abundant schooling species along outer reef dropoffs feeding on zooplankton. The four species not seen at site C presumably would be observed there with further effort. Two of these species are grazers (*Z. flavescens* and *K. bigibbus*) and two feed on benthic invertebrates (*C. aygula* and *P. insularis*). Based on subjective impressions, the habitat availability at site C does not indicate to us any lack of resources or other factors inhospitable to the survival of these species. At the three cluster level, a similar situation arises. When we reach the four cluster level, an interesting pattern emerges. The five sites with northern exposure (sites G, H, I, J, and K) cluster apart from the other sites. This distinction could indicate larval recruitment patterns or environmental factors affect survival along northern shores. Whatever the explanation, this data type reveals definable structure to the communities around the atoll. Notably, the sites surveyed by Lobel and Lobel (2004) were almost entirely on southern exposures which may account for their lack of perceived differences among outer reef habitat types.

The quantitative BLT-ND data type creates an entirely different picture. Again, the clusters are extremely geographically cohesive, but they contain completely different groupings of survey sites. The first split in the dendrogram creates a five-site cluster (sites E, F, G, H, and I) and a seven-site cluster (sites L, J, K, A, B, C, and D). The SIMPER analysis reveals that the dissimilarity between these clusters is due largely to the increased abundance of small bodied planktivores at the five-site cluster. While abundant at all the sites, several species of damselfish (*Chromis acares, C. agilis,* and *C. vanderbilti*) and the anthias (*Pseudanthias pascalus*) are more abundant within this cluster. These five sites are along the western side of the atoll, and none of the five sites are located offshore of emergent land. Mechanisms related to terrigenous input or exposure direction are both plausible explanations for this clustering result. Proceeding along the branches of the dendrogram, the five-site cluster first sheds site E and then site I which are closest to emergent land at the two sites on opposite ends of this cluster.

The BLT-BD data type, while statistically more cohesive than expected at random, seems to have less explanatory power. The first split in the dendrogram isolates sites C, D, E, A and J from the other sites, and these two clusters contain seemingly haphazard sites from all sides of the island. It is not until reaching the 10 cluster level that all clusters from the BLT-BD data type are geographically cohesive. The species identified by the SIMPER analyses as most affecting cluster similarity are primarily mobile, schooling herbivores. A single school of one of these species happening to cross the transect boundaries during a survey could have a drastic effect on the biomass estimate for that site, and there were many such schools in the relatively undisturbed habitat at Wake. Such an event would cause that site to cluster separately from adjacent sites that may be highly similar in all other measures of species biomass, but which differ greatly due to the occurrence of a single high biomass school.

The BLT-PRES data shows the worst geographic cohesion of all the data types, and in fact fails to improve on sites randomly assigned to branch tips. While mathematically robust, the clusters produced from the BLT-PRES data do not lend themselves to a mechanistic explanation of community structure. Given this data type alone, we would feel hard pressed to resolve the structuring of the communities around the atoll.

Two data types showed the least cohesion. The BLT-PRES data was less cohesive than expected from the randomly generated dendrograms, and the BLT-BD data was barely more significantly cohesive than expected. This result is striking since all the data types represent real information collected from the survey sites. Our interpretation is that these two data types are less than ideal in evaluating the relationships between the sites. These two data types suffer from similar methodological problems. They both give greater weight to rare and large species. Certain difficulties in accurately surveying coral reef fishes become apparent from this investigation. The BLT-PRES data is essentially an infinite-root transformation of the BLT-ND data. What is lost in this transformation is that a small number of fish species makes up the vast majority of fish at each site. Contrary to the findings of Meekan and Choat (1997) who observed that the most abundant herbivorous species observed along shallow temperate reefs of New Zealand occurred in mobile schools, the most abundant species at Wake were highly site-attached small schooling planktivores. The variations in the density of these few species are what distinguish sites when using the BLT-ND data, whereas the variation in the rarer species is what distinguishes the sites when using the BLT-PRES data. Some coral reef fish species are extremely cryptic, and the likelihood of accurately documenting the presence or absence of rare species is considerably less than accurately documenting the absence of common species. The limited area of search included within the confined area of the belt transects further exacerbates this problem. As the search area is expanded to include the entire area seen by the diver at the site (RDS-PRES data) the strength of the data grows. Table 2 shows the efficiency of the BLT-PRES data in comparison to the more thorough RDS-PRES data. There is an almost 70% increase in the number of species recorded during the RDS portion of the survey. The BLT-BD data similarly overweights rare species – in this case because large fish tend to be rarer than small fish. While a cloud of 1000 anthias may numerically dominate a site, a single large shark will far outweigh these anthias, and while sharks were seen at nearly all sites, it is less common for them to swim within the boundaries of a belt transect. That said, we do wish to point out the potential importance of investigating biomass relationships between sites. Studies like Gust et al. (2002) have demonstrated nicely that even within single species, growth rates, and subsequently size distributions, can vary between habitats on a reef which would obviously alter relationships of relative biomass between habitats. These habitat differences would be of great interest to scientists and managers attempting to classify communities as effective conservation must take into account habitat usage across all life stages of the species of conservation interest.

As an additional measure of precaution against the possibility of data type sensitivity to some of the technical aspects of the cluster analysis procedures, the analysis was also run using the "group average" clustering mode in Primer. The topologies of the resulting dendrograms was slightly different using this methodology, but the resulting relative counts of cohesion of the data types was nearly identical to the results presented here using the complete linkage clustering mode. The only exception being that the BLT-PRES data became slightly better in terms of the number of cohesive clusters on its tree.

In defining a cohesive cluster, we had to decide how to treat a cluster consisting of a single site. A single site cluster indicates the community surveyed at that location is sufficiently different to warrant its isolation in the cluster hierarchy. This could reasonably indicate insufficient sampling in the area around that single site. It would be expected that additional survey sites in close proximity to the single site cluster would produce a cohesive cluster of multiple sites. Unfortunately, at this remote location we were severely limited in the number of sites we could survey in a relatively short period of time. Identifying the single site clusters could aid investigators in prioritizing further survey efforts. If the investigator's objective is to identify and classify communities within an ecosystem, it would make better sense to add additional survey sites near habitats that have been identified as unusual rather than to continue to add sites in homogeneous areas that have been more adequately sampled.

This investigation revealed several examples of two cohesive clusters joining to form one non-cohesive cluster towards the base of the tree. If we examine the cluster results from the BLT-ND data, we see the two cohesive clusters on the northeast side (sites J and K) and the southeast side (sites A and B) being separated by site L near the southeast tip of the island. It is not unexpected that site L, being situated near a prominent point on the island, is different from the two adjacent clusters. However, it is not expected that the two surrounding clusters, with very different orientations and exposures, would be more closely related to each other than to the site situated between them. The question arises, do similar communities repeat themselves? In a small island like Wake, it would be reasonable to expect that all sites are exposed to the same likelihood of settlement of all species and that the factors that affect community structure are predominantly post-settlement. Sammarco and Andrews (1988) showed how local coral recruitment patterns over a small scale can vary as a function of distance from the brooding stock over short time scales. Yet, they argue that over evolutionary time scales, even for species like corals with a relatively short residence time in the larval stage, recruitment is generally even across areas much larger in scale than Wake Atoll. At Wake, energy regime, nutrient supply, and subsequently available habitat are likely factors affecting local survivorship following settlement, and it would be expected that a gradient of these factors would be evident as one traveled around the ring of the island. The isolation of site L between two similar communities, however, is evidence that communities can in fact repeat themselves.

One other interesting observation that arises from the cluster randomizations is that constraining the topologies of the cluster trees affects the possible outcomes in terms of the number of cohesive clusters possible at different levels of clustering. Looking at the RDS-PRES tree, for example, it is impossible at any clustering level to have 0 cohesive clusters. There will always be a single site cluster which by definition is cohesive. Every time a single site cluster is formed in a tree, it raises the number of expected cohesive clusters. One could imagine a cluster topology where a single site is isolated from the larger cluster at every subsequent level of clustering guaranteeing that all clusters are cohesive with the possible exception of the largest cluster at each level. A tree with that structure would make it more difficult to improve upon expected levels of cohesion. In our study, the two data types that showed the greatest significance were the two data types whose tree topologies produced the highest number of expected cohesive clusters (95%ile: BLT-ND: 49; RDS-PRES: 51; 97.5%ile: BLT-ND: 50; RDS-PRES: 53).

The challenges that face wildlife researchers at remote locations are imposing, and often the window of opportunity is brief. Even in easily accessible areas, scientists often are hampered by a lack of time and resources to survey communities as adequately as they would like. In these circumstances, the field personnel need to know the most efficient and discriminating means to resolve community structure. This study suggests that a combination of numerical density and thorough species presence absence data best serve these needs. However, caution should be used in applying only a single type of data to define community structure as the data type can have great impact on the resulting classification.

CHAPTER 4:
REEF FISH POPULATION TEMPORAL STABILITY

Introduction:

In addition to the concerns addressed in the previous chapter, there are many questions around the temporal stability of fish communities and whether it is possible that temporal instability of the make-up of fish communities could make it impossible to characterize the community at a particular site. Due to the remote nature of many of the study sites in this investigation, some locations were visited only on a single visit. These expeditions typically lasted from 3 to 20 days per location. Some concern has been raised about using surveys with such temporally restricted timeframes as definitive pictures of the communities at particular sites (Stimson, personal communication). Studies of fish communities over time have shown a great degree of variability in structure (Sale and Douglas 1984, Syms and Jones 2000). Many factors may affect reef fish community structure including variation in recruitment (Sale et al. 1984), habitat degradation (Done 1992, Chabanet et al. 1997) and recovery following disturbances (Sano et al. 1987). Additionally, there may be variability over various time frames including lunar, diel, and seasonal periods (Rooker and Dennis 1991). While the author readily acknowledges that fish populations will naturally fluctuate over time, every effort was made in this study to minimize the potential effects of this source of error. Several steps were taken to address this issue. First, several locations were revisited and community structure was compared over time and space using these repeat visits. Non-metric multidimensional scaling (MDS) was used to perform this comparison. Second, the families used to perform the regression necessary to test the hypothesis that population density affects likelihood of dispersal were restricted to those families found to be most stable in population structure over time. To identify these families, the mean normalized change in population structure for species in each family was calculated, and only those families showing high consistency were used in the analysis. It is important to note that the collection of islands in this study are all largely uninfluenced by human pressures and can be considered in as natural a state as exists in coral reef systems. Their remoteness, while offering high levels of de facto protection from impacts associated with human fishing and development pressures, also means that visits to them by scientists is very infrequent. As such, there was no readily available information concerning the state of recovery of these island systems from natural disturbances like storms, Acanthaster outbreaks, bleaching events, or others.

Methods:

Fish populations were quantitatively assessed using a modified belt-transect method (BLT) following Friedlander & DeMartini (2002). The author swam along a 25m transect line recording abundance of all fish >20cm TL observed within a 4m wide belt along the transect line. He then returned along the line recording abundance of all fish <20cm TL observed within a 2m wide belt. All fish were identified to the lowest possible taxon which was the species level in nearly all cases. Three transects separated by 5m were performed at each site. All transects followed a constant 15 m depth contour and continued in the same direction regardless of changes in habitat that were encountered. A total of 31 islands were surveyed using these methods. Four of these islands were revisited two years after their initial visit: Howland, Baker, Jarvis and Swains at approximately the same time of year thereby minimizing any possible seasonal fluctuation effects that might otherwise confound data analyses. Non-metric multidimensional scaling (nMDS) was performed using Primer-E and was based on Bray-Curtis similarity indices calculated between all islands based on fourth-root transformed numeric density estimates (Clarke and Warwick 2001). For the four islands surveyed twice, the individual survey years were treated as separate islands. The resulting MDS plots were used to examine the multivariate shift in community structure over time at these four islands and to compare them with levels of similarity between the other islands surveyed in this investigation.

To further isolate the families that would be used in subsequent regression analyses, within-family temporal stability was calculated by taking the mean normalized percent change in population density for every species in each family at the four revisited islands. Upon examination of these results, the families with high variability were conservatively eliminated from subsequent analyses.

Results:

The MDS plots (Fig 4.1) show the same strong relationships between closely situated islands as seen in the cluster analyses performed in chapter 2 based on species presence-absence records at each island. The Northwest Hawaiian Islands form one group, the main Hawaiian Islands form another. Johnston Atoll sits somewhere in between the Hawaiian Islands and the rest of the Pacific, as does Wake Atoll. The Marshall Islands are closely related, and the islands of American Samoa are all very similar too.

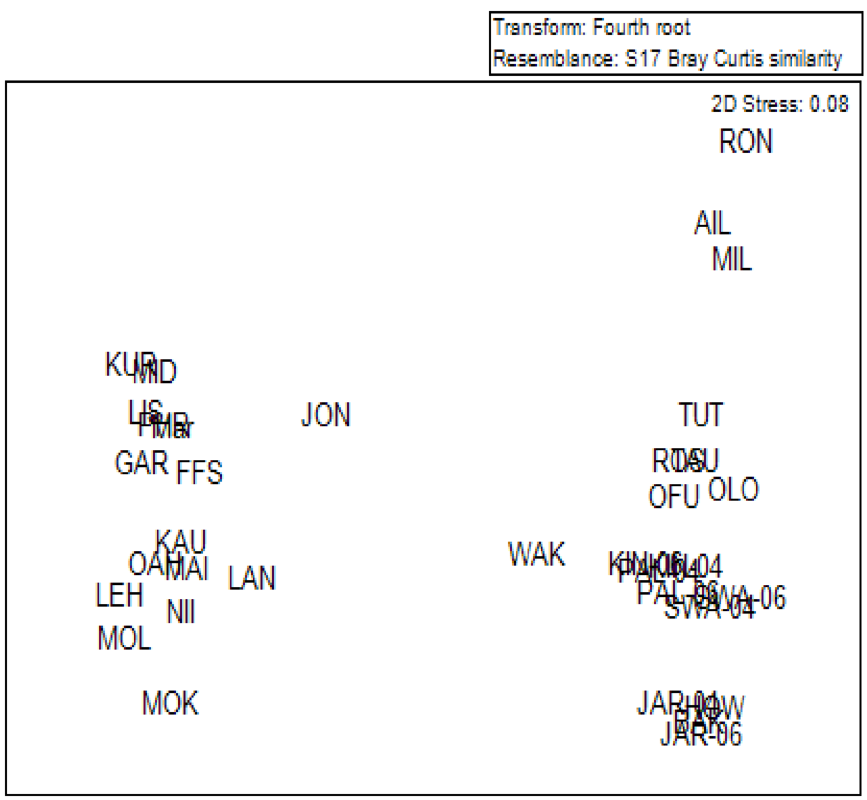

Figure 4.1. nMDS ordination plot of all islands surveyed including the four islands that were sampled twice. Abbreviations follow those listed in chapter 2.

The Line and Phoenix Islands as well as Swains Island must be examined more closely to see the exact relationships between the islands there and are shown in more detail in Figure 4.2. This close up shows a high level of similarity through time for all four resurveyed islands.

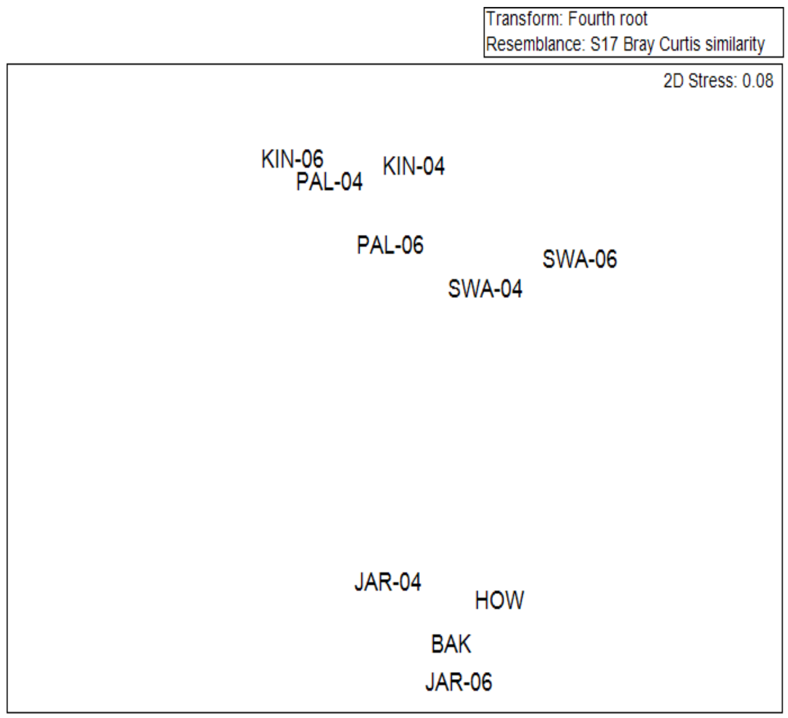

Figure 4.2: Close-up of the US Line Islands, US Phoenix Islands, and Swains Island from Figure 4.1. Close-up was necessary as the highly similar community structure of these islands can not be resolved from Figure 4.1. Acronyms follow the list in Chapter 2 with the addition of survey year for the four resurveyed islands (04=2004, 06=2006).

All of the resampled islands appear to closely resemble themselves through time. Even given a two year gap between sampling times, inter-island differences in community structure appears to be much greater than intra-island differences between temporally separated surveys at the same island. This suggests that while fluctuations do exist in the dataset with respect to population numbers for each species, the scale at which they vary is insignificant when compared with the much larger differences that are inherent regionally.

Within all the families observed, mean normalized change in population density ranged from 0.34 for the single-species Zanclidae family up to 2.19 for the Carangids.

FAMILY	# of Species	Temporal Change
Cirrhitid	7	0.87
Serranid	17	1.10
Lutjanid	8	1.24
Mullid	6	1.30
Chaetodontid	18	1.34
Labrid	46	1.38
Pomacentrid	20	1.46
Acanthurid	31	1.58
Balistid	8	1.62
Holocentrid	7	1.65
Scarid	14	1.68
Carangid	7	2.19

Table 4.1: List of all families with more than 5 species observed along with the mean temporal change in numeric density for all species within each family.

Given the constraints of the model to be used in the test of the hypothesis that species density may be a factor in determining the likelihood of long distance dispersal, only families with more than five species were considered as suitable families to include in the model so as to insure sufficient degrees of freedom for the necessary regression model. Among these families, only those scoring in the top 60% of this test of temporal stability were included in subsequent analyses. It should be noted that there is no absolute test for stability through time that can be conducted to say with certainty that a population is stable enough to be included in the subsequent regression. These measures to limit the families were established arbitrarily to use only the most conservative group of fish families that are presumably more stable over time and that are most accurately surveyed using visual census methods. Following the application of these two measures, the families that remained were the following: Cirrhitidae, Serranidae, Lutjanidae, Mullidae, Chaetodontidae, Labridae, and Pomacentridae. Subsequent treatment and discussion of their inclusion is discussed in Chapter 5.

Discussion:

The main investigation of this dissertation centers around the question of whether population density at a source location affects likelihood of dispersal across large oceanic distances. Experimental manipulation on the scale necessary to directly test this hypothesis is not practical or possible. One of the assumptions of the model detailed in Chapter 5 is that the population densities measured in this study reflect temporally stable enough populations over long periods of time to have meaning over time scales that far exceed the ability of human observation to measure. It is fair and reasonable to suggest that periodic disturbance events might cause temporary or even permanent shifts in the abundances of species at source locations (Sale et al. 1984; Sano et al. 1987; Done 1992; Chabanet et al. 1997), and that during these times of shift, the likelihood of a temporarily abundant species successfully dispersing might increase (if the central hypothesis of this dissertation holds true). While the author acknowledges this potential shortcoming of the approach taken to investigate this question, this chapter represents the best possible precaution available given the scope of time with which this investigation could take place. It is worth suggesting that this investigation could simply serve as the starting point of ongoing, longer term investigations similar in nature to see if any significant relationships hold true over longer time scales.

One notable factor in the islands that are used in this analysis is their lack of influence of human development and fishing pressures. All of the source islands included in this analysis are unpopulated with the exception of Swain's which has a population of roughly 30 people and Palmyra which has a small scientific research station which houses a small number of resident scientists. Presumably, the snapshot in time which the author was able to collect in regards to the population densities of fish species at these islands reflect islands in a very natural state free from the density altering impacts that humans often bring.

CHAPTER 5:

THE RELATIONSHIP BETWEEN SOURCE POPULATION ABUNDANCE AND THE LIKELIHOOD OF LONG DISTANCE DISPERSAL IN CORAL REEF FISHES

Introduction:

Biogeographers have long been attempting to explain why certain species exhibit wider distributions than others (Guisan and Thuiller 2005). In coral reef fish systems, where long distance movements of adults is rare, much attention has been paid to the idea that pelagic larval duration (PLD) is the determining factor in a species' ability to achieve a widespread distribution (Lester and Ruttenberg 2005; Shanks 2009). The open water between islands probably does not offer the type of food normally consumed by adult reef fishes and the obvious lack of shelter doubtless results in heavy predation risk. By contrast, the larval stages of many reef species are often well suited for extended periods of time in the pelagic realm and it has been argued that it is larval dispersal that allows a species to cross the large distances that often separate islands that offer suitable habitat for the adults. If the pelagic larval stages of a species are able to remain adrift for a longer period of time, the stochastic forces of wind and currents which largely drive the passive dispersal of these larvae will be more likely to bring a given larvae close enough to a potential home where the larvae will be able to settle out and successfully grow into a reproductive adult at the new ("destination") location.

Different species have shown different propensities for the length of time which they are capable of surviving in their larval stage but tests of the relationship between PLD and dispersal have shown mixed results (Lester and Ruttenberg 2005). While some studies have shown correlations between PLD and species range, these correlations are often very weak if they are detectable at all (Victor and Wellington 2000; Zapata and Herrón 2002). At the very large end of the range size, a threshhold PLD of 45 days has been used to suggest a larval duration at which widespread distribution would be expected (Brothers and Thresher 1985). However, within the middle ranges of PLD, the relationship between larval duration and range size or genetic connectivity is not observed (Lester et al. 2007; Weersing and Toonen 2009). Additionally, PLD has been observed to be dynamic even within a single species depending upon time of year and by location where a species is sampled (Bay et al. 2006). The fact that PLD has been seen to be dynamic and dependent on location further calls into question some of the conclusions of the previous studies. While some studies were able to show a relationship between PLD and distribution, the fact remains that these results are correlations – and correlation does not necessarily equate to causation. It is possible that these species that are sampled at remote and isolated locations are showing longer PLDs simply because they stayed in the water column longer due to the isolation of the location at which they were able to settle. Perhaps isolation of a location is driving lengthening of PLD locally instead of PLD determining which species are able to cross long barriers. At any rate, there are clearly alternative mechanisms at work in determining a species geographic distribution.

One possible alternative is that the quantity of larval production at a source location may influence likelihood of long distance dispersal. Likelihood of a species' offspring reaching a distant location could be expected to be very low when the distance between sites is far. While a model that considers duration in the water column is surely logical, the next reasonable idea to test would be this: how would the quantity of larvae produced at the source location affect this likelihood. It seems reasonable to expect that if a parent population creates a very small number of offspring, the likelihood of dispersal would be lower than if a population produces a huge number of offspring. The mixing that occurs over time during the dispersal process surely would favor the population that has expelled a larger number of larvae over the population that has only released a handful of possible settlers. And while this hypothesis seems logical, the practicalities of working in these systems have made testing it as an hypothesis impossible to date. Directly measuring the quantity of larval production of a species on the scale of an island or a region is not feasible. As such, a suitable proxy for larval output had to be utilized. The metric chosen here is the numeric abundance of adult individuals of a single species. Following on the work described in the previous chapters, this hypothesis is tested here.

Methods:

Reef fish populations were surveyed at 10 islands in 5 regions using modified belt-transect methods described in chapters 2 and 3. The regions were as follows:

Region	**Region Acronym**	**Islands**	**Distance to Johnston (km)**
Johnston	JON	Johnston Atoll	N/A
Wake	WAKE	Wake Atoll	2,550
Marshall Islands	RMI	Ailinginae, Rongelap, Mili	2,350
Line Islands	LINE	Jarvis, Kingman, Palmyra	1,365
Phoenix Islands	PHOEN	Howland, Baker	1,930

Table 5.1: Islands and regions included in the multiple logistic regressions along with the acronyms used in the results section and the distance measured in kilometers from Johnston Atoll to the closest island of each region.

Within each region, population density for each species was summed across all transects performed. No correction was made for either the number of transects performed in each region or the available habitat size of each island in each region as these corrections would only affect resulting regression coefficients and not the existence of significance in the regression relationships. The relevant question is simply the existence of significant relationships between population density and dispersal, and coefficients of these relationships were not considered to have strong biological meaning.

While the hypothesis being tested is that abundance of larvae produced into the water column may affect likelihood of dispersal, propagule abundance was not directly measured. This would be impractical and logistically nearly impossible. Instead, the numeric density of reproductive adults was used as a proxy for propagule abundance. There is considerable variation in fecundity between coral reef fishes, so to minimize this effect, tests were performed separately for each family included in the analysis.

Additionally, fecundity within a species is known to be related to the length of an individual in an exponential manner, with a general relationship being cubic but with relationships up to the fifth power being recorded (Jennings et al. 2001). While it was the original intention to use biomass of individuals present at a source location for these regression analyses (biomass is also related approximately cubically to fish length), based on the results from Chapter 3 of this investigation, biomass approximations from the employed visual survey methods was deemed unsatisfactory.

Johnston Atoll was considered the target for colonization, and a model was built to test the hypothesis that the abundance of a given species at one of the four potential source regions would influence the likelihood of long-distance dispersal to Johnston. These four regions create an arc around Johnston and represent all of the closest islands to Johnston (Fig. 1). Lobel (1997) suggested that the fauna of Johnston may be tightly linked to that of the Line Islands but are probably predominately self-seeding (Lobel 1997). Lobel described the fish fauna of Johnston as a mixture of Hawaiian island species with Line Island species. While this is partially accurate, as can be seen from the dendrogram in chapter 2, there are other regions – principally the Phoenix Islands, the Marshall Islands, and Wake Atoll which also share some commonality with Johnston in terms of the fish fauna present.

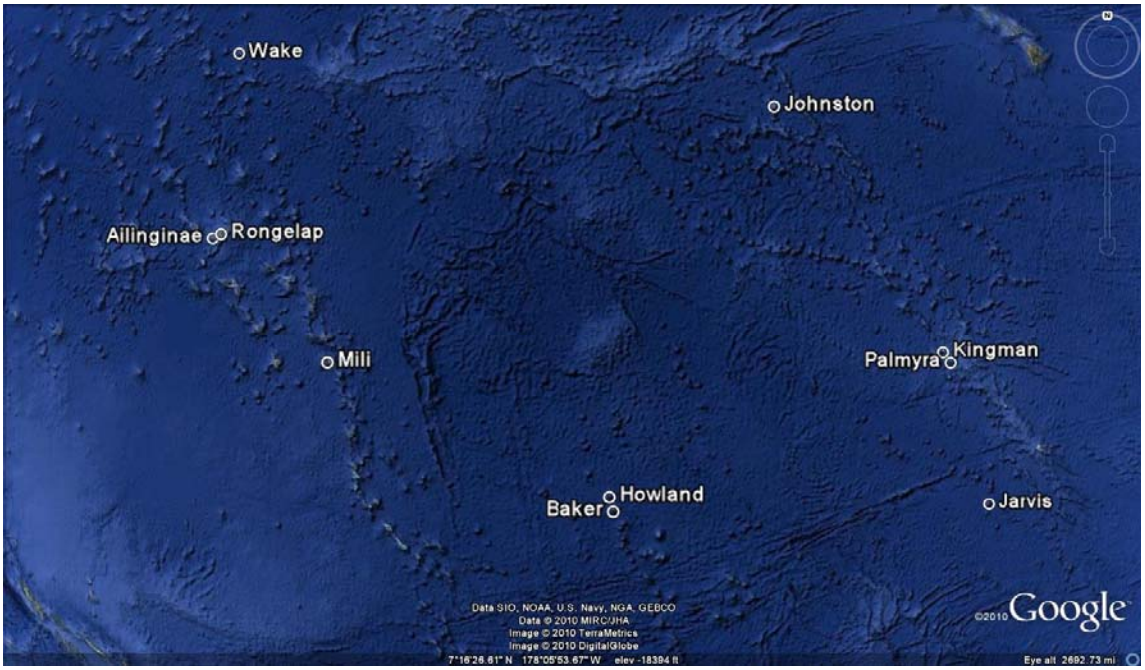

Figure 5.1: Map of the study region showing the islands situated in a ring around the destination location of Johnston Atoll.

Based on the observations during the roving diver surveys, a comprehensive species list was compiled for all islands, and the list for Johnston Atoll was supplemented with literature records of known species from the region that were not recorded during this survey (Kosaki et al.; Randall et al. 1985; Gosline 1955). A species' presence or absence at Johnston was used as the response variable *z* in the multivariate logistic regression model:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k,$$

where β_k is the regression coefficient for species *k*, and χ_n is numeric density at region *n* as estimated from the belt-transect surveys conducted in each region. Models were run in the statistical software package R (R Development Core Team, 2004) and were tested through Wald Chi-Square Tests of maximum likelihood estimates with a threshold significance of 0.05 using the "lrm" function in the RMS package. Models were built for the species of each family with terms representing each region. The least significant regions were dropped sequentially until all remaining terms in the model were considered significant. Models were tested for goodness of fit using the Hosmer-Lemeshow-Cressie test (Hosmer et al. 1997) from the residuals.lrm function in the Design package. For single term models, plots of density versus dispersal success were generated and lowess lines were added to visually assess the model fits.

Results:

Based on the tests of temporal stability in Chapter 4, seven families were identified as sufficiently speciose and sufficiently temporally stable to be included in this analysis: Cirrhitidae, Serranidae, Lutjanidae, Mullidae, Chaetodontidae, Labridae, and Pomacentridae. Two of these families could not be tested. The Serranidae were not tested because none of the species present in any of the four source regions are present at Johnston. This precludes the mathematics of the model from producing meaningful results. Similarly, the Mullidae were eliminated because all of the species recorded on quantitative transects at the four source regions were also recorded at Johnston. The remaining five families were tested, and of them, three families showed significant results with population density at a source location being directly correlated with the likelihood of occurrence at Johnston. The remaining terms in the best-fit models along with their p-values are shown in Table 5.2.

Family	Region	p-value	Model AIC
Chaetodontidae	Line Islands	0.03325	44.959
Cirrhitidae	n.s.		
Labridae	Line Islands	0.00585	67.605
	Phoenix Islands	0.03188	
	Wake Atoll	0.02855	
Lutjanidae	n.s.		
Pomacentridae	Wake Atoll	0.0301	31.85

Table 5.2: Results of logistic regression models showing significant terms remaining in best fit models. P-values for each term included in the final models are included along with the overall AIC for each model. Families where no significant relationship was

found are indicated by n.s.

Based on the Hosmer-Lemeshow-Cressie goodness of fit tests, all three models failed to show any significant lack of fit (Chaetodontidae p=0.567; Labridae p=0.361; Pomacentridae p=0.443). Plots of the Chaetodontidae and Pomacentridae models (figure 5.2) show good fit of the expected logistic regression relationships for both models with the expected sigmoid shape of a logistic regression model seen clearly in the lowess smoother curve for both models.

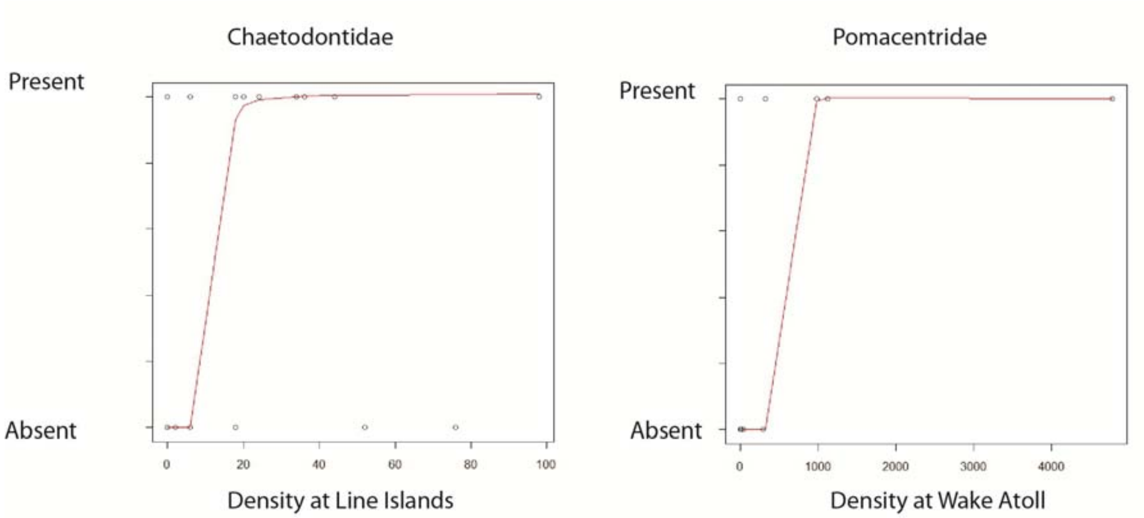

Figure 5.2. Presence or absence at Johnston Atoll versus density at significant source locations for the Chaetodontidae and Pomacentridae families. Red line is a lowess smoother that shows the fitted relationship of the data.

Discussion:

Based on the assumption that adult population density can be used as a reasonable proxy for larval propagule abundance, the results presented here provide strong support for the hypothesis that high source location larval output increases the likelihood of long distance dispersal in coral reef fishes. While two tested families showed no significant relationship between adult density at putative source locations and occurrence at Johnston Atoll, three out of five did.

Two of the regions in particular showed up as significant factors in multiple families – Wake Atoll and the Line Islands. Based solely on geographic distance to Johnston, these results are somewhat unexpected as Wake is the furthest of all the regions from Johnston (Table 5.1). The Line Islands, however, are the closest region to Johnston, so based solely on geographic distance, their significance in the models could have been predicted to be higher. Wake (19 deg. North) is, however, situated at a similar latitude as Johnston (16 deg. North) even though the northern Atolls of the Marshall Islands region (11 deg. North) are also close in latitude and much closer in terms of distance.

An examination of prevailing currents in the region offers some additional explanatory power in evaluating why Wake Atoll was included in two of the three models. Current models show a strong connection between Wake Atoll and Johnston due to the Eastward flowing subtropical countercurrent and the Hawaiian lee countercurrent (Kobayashi 2006) which could help explain the connection between Wake and the Johnston/Hawaii system, but similar current patterns do not exist connecting the Line Islands region to Johnston Atoll. One important possibility to consider is the impact that storms and other irregular current and wind events could play in the ongoing dispersal between regions. If regular current patterns do not favor transport to Johnston from the Line Islands, but the occasional storm pattern disrupts the normal current regime, then even further credence to the hypothesis of propagule abundance affecting dispersal likelihood might be warranted. During a highly turbulent storm-type event, the benefits of having a large number of larvae set adrift in the water column would be magnified and may increase in importance relative to the pelagic larval duration of a given species.

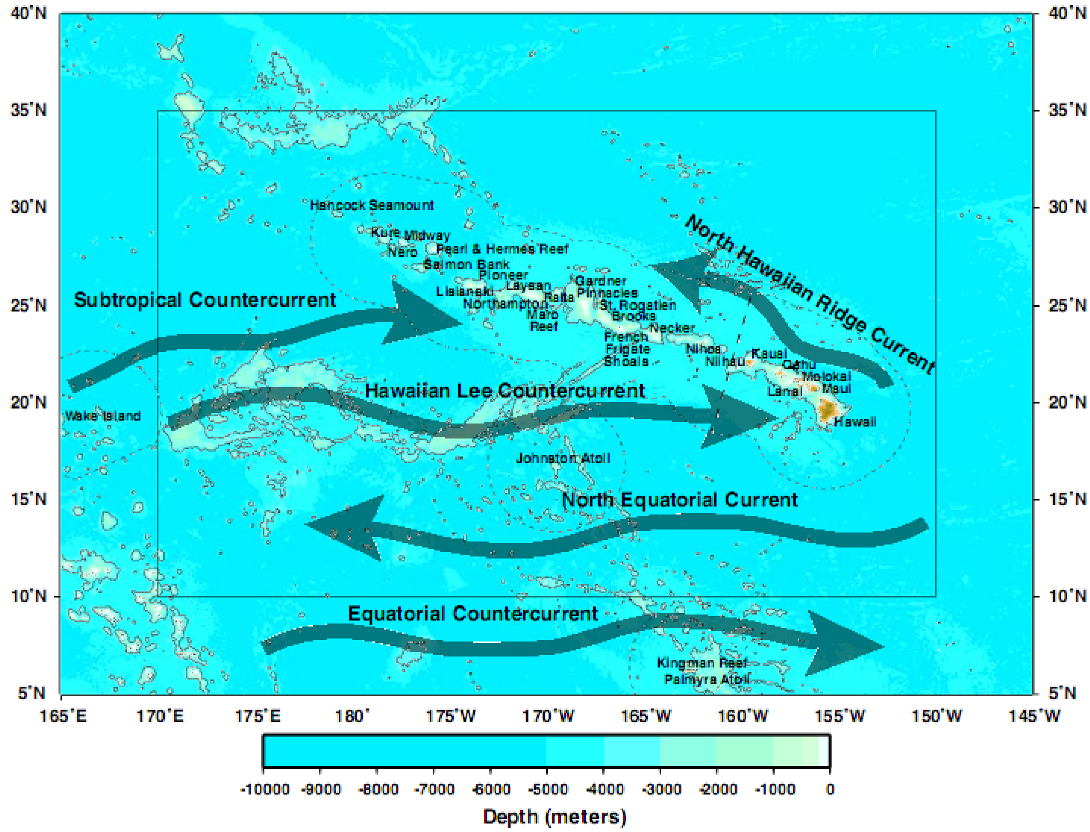

Figure 5.3 Map of Johnston Atoll, the Hawaiian Island chain, Wake Atoll and the Line Islands with dominant current systems. Reproduced from Kobayashi 2006.

If population density affects dispersal likelihood, the question arises as to what factors affect population density at source ("upstream") locations. Are there ecological factors at work throughout the region that cause certain species to be less abundant at an upstream location? And are these factors responsible for the inability of certain species to produce enough offspring that ongoing long distance dispersal is sufficient to maintain a species across large barriers?

There are certainly examples of coral ref fish that have very specific habitat requirements. Anenomefish, for example, require certain species of host anenomes to exist. And while there are also many generalist types of species with far less narrow habitat requirements, it is probable that habitat and environmental conditions can have an impact on a given species' population density. While we are unable to say with certainty that a species' lack of presence at a destination location (like Johnston) necessarily means that there has been a failure of dispersal to that location (rather than a failure to colonize at that site after arrival), for the purposes of this study, we assume this is the case.

Experimental manipulation on the scale of this study is impossible. It is not feasible to manipulate source location population density to see if it has an effect on dispersal success, and as such we will be largely left with results from studies like this one as the best basis from which to draw conclusions. However, there has been one interesting action taken that may provide some insight into this situation. Beginning in the 1950s, the introduction of several species of fish to Hawaiian waters was undertaken by the State department of Land and Natural Resources. Among the introduced species, three species in particular – *Cephalopholis argus, Lutjanus kasmira,* and *Lutjanus fulvus* – achieved a high level of reproductive success in Hawaii, and a fourth species – *Lutjanus gibbus* – is occasionally observed by divers and fishermen in Hawaiian waters as well (Schumacher, personal communication). And while the level of ecological success has differed between these species, all have successfully established themselves in Hawaii. These three species all have broad distributions around the Pacific, yet none had ever successfully established breeding populations at Johnston Atoll or in Hawaii. Since their introduction, they have proliferated throughout the Hawaiian island chain with reports of presence on almost all the islands including the remote Northwest Hawaiian Islands. The most successful of these species – *Lutjanus kasmira* – has been reported as far north as Midway Atoll – although it is not yet known from Kure, the northernmost atoll in the archipelago (Schumacher & Kosaki, personal communication). These species' ability to successfully survive in the Hawaiian Island region clearly suggests their range was limited by dispersal ability and not by ecological factors at the destination location. Interestingly, these species are known from all of the source locations used in this study. However, they were recorded at very low population densities at the sources. Unfortunately, the Serranidae family was not able to be tested in this study, and the Lutjanids were tested with certain caveats discussed below.

Five families were tested in this investigation (Chaetodontidae, Cirrhitidae, Labridae, Lutjanidae and Pomacentridae). These families represent a wide variety of fish species with respect to reproductive strategies, social systems, and general ecological characteristics. All are typically small bodied (with a few notable Labrid and Lutjanid exceptions) and well suited to being accurately counted by *in situ* underwater visual census methods. While the Pomacentrids are predominately demersal egg laying species, the other four families are pelagic spawners.

The two families that showed no relationship between source location population density and dispersal success (Cirrhitidae and Lutjanidae) deserve some attention. The Lutjanidae were included in the analysis even though only two of the species enumerated at the source locations were known from Johnston. These two species, *Aprion virescens* and *Aphareus furca*, are significantly different from the other members of the Lutjanidae family – the rest are of the genus Lutjanus. While a comparison of the fecundity of these two species to the rest of the species in the Lutjanidae family is not reported in the literature, their ecology, morphology, and taxonomy are extremely different from the other members of the family. Consequently, the inclusion of the Lutjanidae family in the analysis was questionable as testing them may not have directly addressed the relevant question at hand. The potential confounding factors surrounding the dissimilarity of these two dispersive species versus the members of the genus Lutjanus indicates that analysis of the Lutjanids should not be given as much weight as the other families.

The other family (Cirrhitidae) that showed no significant relationship between population density and dispersal success for any of the regions is typified by species that are highly site-attached and not very mobile (Donaldson 1990). The remaining three families that constitute the core of this analysis generally consist of species that are more mobile – although certainly examples exist of highly territorial Pomacentrids with very small home ranges on the order of a few square meters or less (Sale 1971; Ross 1978). Unlike the Lutjanids (for which there are compelling reasons not to include them), no such disqualifying attribute can be found among the Cirrhitid species. It must therefore be assumed that there are other unknown factors in the biology of the Cirrhitidae that account for the lack of dispersal.

Reproductive systems among the tested families include a whole suite of strategies ranging from lifelong pair bonders (certain species of Chaetodontids) (Hourigan 1989) to haremic dominant males (certain Labrids) (Donaldson 1995) to protandrous hermaphrodites (some Pomacentrids) (Ross 1978b). Such a wide variety of ecological characteristics in these three families suggests that the influence of population density on larval dispersal ability seen may hold true for many types of reef fish and has broad general applicability.

Conservation Ramifications

Since local population density is affected by ecological and environmental conditions (Hixon & Beets 1993), then the results of this study may have broad implications in conservation. As conditions change that alter population structure at a source location, species that become increasingly rare may exhibit reduced gene flow to destination locations. While the current paradigm of reef fish replenishment heavily favors locally spawned individuals resettling on their home reefs (Jones et al. 2006), there is clearly some level of gene flow between widely disparate locations that maintains intra-species genetic complexity across these large boundaries (Weersing and Toonen 2009). If this dispersal is reduced or eliminated, reduced genetic diversity at a destination location could be realized, and in extreme cases, complete isolation could occur. Thus, it seems this study provides additional support for globally integrated conservation strategies that consider implications far beyond the shores of a locally managed area as population density has been shown to affect dispersal likelihood.

A Posteriori Presence/Absence Analysis

While it is the author's belief that the existence of a density effect on the likelihood of dispersal from the source regions to Johnston Atoll was well demonstrated in three of the five families tested, one analysis was suggested as an a posteriori alternative hypothesis and the results of that test are included here (Taylor, personal communication). Upon examination of the collected data, it was hypothesized that the relationship seen between source population density and dispersal likelihood might be equally well described by examining simply the relationship between presence/absence of a species at the source location and dispersal success. The strength of the significant relationships seen in the analysis of the density models seemed to be largely driven by a high number of low or near zero density species at the source locations that were also not found at Johnston Atoll. To test the strength of presence/absence models relative to the density models, logistic regressions were repeated using only presence/absence at the source locations and ignoring the additional information of population density. Results from these analyses are shown in table 5.3.

Family	Region	p-value	Model AIC
Chaetodontidae	Line Islands	0.004467	41.718
Cirrhitidae	n.s.		
Labridae	Wake Atoll	0.00000632	71.097
Lutjanidae	n.s.		
Pomacentridae	Phoenix Islands	0.0473	47.385

Table 5.3: Results of logistic regression models based on presence/absence analyses showing significant terms remaining in best fit models. P-values for each term included in the final models are included along with the overall AIC for each model. Families where no significant relationship was found are indicated by n.s.

The results of these analyses were then compared with the results from the analyses that included population density (Table 5.2). Several interesting results are seen. First, again the Cirrhitidae and Lutjanidae families show no significant relationship from any source region. The Chaetodontidae family again reveals the Line Islands region to be a significantly related source, and while the Labridae and Pomacentridae families also show significant relationships, the regions which are identified as significant are different from the regions identified in the density models. When the models for each family are compared using the AIC method, only the Chaetodontidae family model shows an improvement using the presence/absence model over the density model as the AIC for the presence/absence model (41.718) is lower than that of the density model (44.959).

Based on the presence/absence model, the Labridae family only shows a significant relationship for the Wake Atoll region, while the density model includes Wake, the Line Islands, and the Phoenix Islands. Based on comparisons of AIC, the density model (67.605) is a better fit than the presence/absence model (71.097).

For the Pomacentridae family, the presence/absence model shows a significant relationship for the Phoenix Island region, but again the AIC of this model (47.385) is higher than that for the density model (31.85) that shows a significant relationship for the Wake Atoll region.

While one of the presence/absence models does show an improvement over the density models, there remains strong evidence of a density effect in at least two of the tested families. When examining the density of species observed in the three families that show significant relationships, it is interesting to note that there is a great disparity in the range of density estimates for the species in these three families. The Chaetodontidae family is made up of species with extremely low recorded densities when compared to the Labridae and the Pomacentridae. For the Chaetodontidae species observed at the Line Islands region, the mean density of all species was 33.4 individuals and ranged from 2 to 98. For the Pomacentridae, this mean value was 2193.4 individuals and ranged from 2 to 21,820. For the Labridae, the mean value was 168.4 and ranged from 1 to 2048. It is possible that the extremely low densities of species in the Chaetodontidae family allow presence/absence to be very nearly as good or better predictor of dispersal likelihood than population density.

CHAPTER 6:

SUMMARY

In the preceding chapters several steps were taken to address the question: "*Does the density of adults in a coral reef fish population at a source location affect the likelihood of that population being able to successfully disperse across a broad geographic distance and colonize remote areas?"*

A large-scale data collection effort was undertaken to assess the population density of hundreds of species of coral reef fish across dozens of islands in the Central Pacific. A total of 772 fish species were recorded at 38 islands. Over 700 belt-transect surveys were conducted which enumerated over 300,000 individual fish. Such a broad-scale study with a standardized quantitative methodology conducted by a single individual is unique, and it allows for rigorous testing of a novel and heretofore untested hypothesis concerning the mechanisms of reef fish dispersal.

The first step in this investigation involved analyzing the visual census data to identify any regional groupings of islands. Multivariate cluster analyses based on the observed presence and absence of species at each of the 38 study islands was used to elucidate these relationships. This method revealed strong spatial autocorrelation (as expected) and defined distinct groups of islands that could be used as the basis of regional groups from which to test the central hypothesis of this dissertation. The data supported the already well documented major biogeographic gap between the Hawaiian Islands and the rest of the Pacific islands and revealed Johnston Atoll as a probable 'stepping stone' or 'gateway' for species expanding their range from Western and Southern islands to the remote Hawaiian archipelago. This initial analytical procedure set the framework for testing the major hypothesis addressed in this study and also served as the preliminary validation of the sampling design and *in situ* visual census data collection protocols for accurately assessing the reef fish communities in these locations.

To further address methodological concerns regarding the utility of *in situ* visual census methods, several methods of fish population assessment were rigorously examined. This examination revealed shortcomings in some of these methods and resulted in the development and implementation of new statistical methods to insure that only the most robust and valid data types were used in the ultimate regression analyses. The technique that was developed involved testing the ability of each of four data types to reveal faunal relationships that reflected the spatial relationships between the study sites. This analysis was based upon finer scale analysis of data collected at a single study site – Wake Atoll. Because the study sites at Wake Atoll were situated in a ring around the atoll, special consideration was necessary in the technical development of the census methodology and the resulting method was termed “circular ordination.”

Four data types were tested for their robustness. Ultimately, species checklists from roving diver surveys and numeric density from belt transects proved to be more robust than biomass estimates and presence-absence data limited to the species observed within the limited dimensions of belt transects. Multivariate cluster analyses of the two most robust data types were able to produce clusters of sites that were more cohesive at all levels of relatedness along the branches of the resulting dendrograms than randomly generated clusters of sites. These analyses had an extremely high level of statistical significance. This validation of the two most powerful data types supported the use of these types of data in the other sections of this investigation.

To further protect against the possibility of using less than sufficiently reliable data in the final hypothesis test of this dissertation, an additional measure was taken to test for and minimize the impact of temporal instability in the census data. Temporal stability of fish population data was tested as rigorously as possible given the limited opportunities to revisit the remote and difficult to access study locations included in this study. First, non-metric multi-dimensional scaling plots were generated for all sites where quantitative species abundance was recorded, with revisits to islands treated as distinct study sites. The resulting plot showed clearly how the differences between islands was substantially greater than the differences between subsequent visits to the same island for all four islands which were visited on a second occasion. While no specific threshold level of temporal stability could be used as a basis for inclusion of data in the logistic regression of the final chapter, data from the limited number of islands that were visited on multiple occasions were compared for temporal stability at the species level within families. An operational decision was made that data used in testing the colonization hypotheses would be restricted to the top 60% of fish families in terms of stability exhibited over successive sample periods. To calculate temporal stability, the mean standardized change in species density was calculated for each fish family. Ultimately, these temporal stability precautions reduced the number of suitable families for testing down to seven families. Unfortunately, two of these seven families could not be used because of insufficient numbers of both binary response variables (absent or present at Johnston).

Finally, a series of multiple logistic regression analyses were performed to address the central hypothesis of the investigation: population density at a source location affects the likelihood of long distance dispersal. The regressions were based on the numeric density data which were identified as a statistically robust in Chapter 3 and were limited to the families identified as the most temporally stable in Chapter 4. The resulting regressions showed strong statistically significant relationships between populations at Johnston Atoll and populations at other locations as follows:

Family	Region
Chaetodontidae	Line Islands
Labridae	Line Islands, Phoenix Islands and Wake Atoll
Pomacentridae	Wake Atoll

No significant relationships were found for the Cirrhitidae or Lutjanidae families from any region. The Marshall Islands region was not an important source for Johnston Atoll for any of the families tested.

The results show that that the sheer size of a reef fish ‘source’ population (rather than geographical proximity or other biological factors such as pelagic larva longevity) is an important factor in whether species can successfully colonize remote areas. The implications of this study are broad. First, if ecological conditions at ‘upstream’ locations are impacting the local abundances of fish species, they could in turn be affecting the chances of long range dispersal. While this has meaning in an evolutionary sense and over time scales beyond the scope of direct human investigation, there is also implication in human terms. As conditions continue to change due to human impacts, long range gene flow to ‘downstream’ locations could also be impacted. If species densities are reduced below threshold levels, genetic isolation could follow.

BIBLIOGRAPHY

Bay, L. K, K. Buechler, M. Gagliano, and M. J Caley. 2006. Intraspecific variation in the pelagic larval duration of tropical reef fishes. *Journal of Fish Biology* 68, no. 4 (April 1): 1206-1214.

Bonhomme, F. and S. Planes. 2000. Some evolutionary arguments about what maintains the pelagic interval in reef fishes. *Environmental Biology of Fishes*. 59: 365–383.

Booth DJ, Beretta GA (2002) Changes in a fish assemblage after a coral bleaching event. *Mar Ecol Prog Ser* 245: 205-212

Brock, R.E., 1982. A Critique of the Visual Census Method for Assessing Coral Reef Fish Populations. *Bulletin of Marine Science*, 32(1), pp.269-276.

Brothers, E. B, and R. E Thresher. 1985. Pelagic duration, dispersal and the distribution of Indo-Pacific coral reef fishes. *The ecology of coral reefs* 3: 53–70.

Chabanet, P., H. Ralambondrainy, M. Amanieu, G. Faure, and R. Galzin. 1997. Relationships between coral reef substrata and fish. *Coral Reefs* 16, no. 2 (June): 93-102.

Clarke, K.R. and Warwick, R.M. 2001. Change in marine communities: an approach to statistical analysis and interpretation, 2nd edition. PRIMER-E, Plymouth.

Darwin, Charles. 1872. *On the Origin of Species by Means of Natural Selection, or, The Preservation of Favored Races in the Struggle for Life*. 5th ed. New York: D. Appleton.

Dieckmann, U. & Doebeli, M., 1999. On the origin of species by sympatric speciation. *Nature*, 400(6742), pp.354–357.

Donaldson, T.J., 1990. Reproductive behavior and social organization of some Pacific hawkfishes (Cirrhitidae). *Ichthyological Research*, 36(4), pp.439–458.

Donaldson, T.J., 1995. Courtship and spawning of nine species of wrasses (Labridae) from the western Pacific. *Japanese Journal of Ichthyology*, 42, pp.311–320.

Done, T. J. 1992. Phase shifts in coral reef communities and their ecological significance. *Hydrobiologia* 247, no. 1-3 (November): 121-132.

Fowler H, Ball S (1925) Fishes of Hawaii, Johnston Island, and Wake Island. Ber P Bish Mus Bul 26: 1-31

Friedlander AM, DeMartini EE (2002) Contrasts in density, size, and biomass of reef fishes between the northwestern and the main Hawaiian islands: the effects of fishing down apex predators. Mar Ecol Prog Ser 230: 253-264

Gosline, W.A., 1955. The Inshore Fish Fauna of Johnston Island, a Central Pacific Atoll. Pac Sci 9(4): 442-480.

Guisan, A. & Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8(9), pp.993–1009.

Gust N, Choat JH, Ackerman JL (2002) Demographic plasticity in tropical reef fishes. Mar Biol 140: 1039-1051

Hixon, M.A. & Beets, J.P., 1993. Predation, Prey Refuges, and the Structure of Coral-Reef Fish Assemblages. *Ecological Monographs*, 63(1), pp.77–101.

Hosmer, D.W., Hosmer, T., Le Cessie, S. and Lemeshow, S. 1997. A comparison of goodness-of-fit tests for the logistic regression model. Statistics in Medicine. Vol. 16. 965-980.

Hourigan, T.F., 1989. Environmental determinants of butterflyfish social systems. *Environmental Biology of Fishes*, 25(1), pp.61–78.

Hubbell, Stephen P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton University Press, May 1.

Humboldt, Alexander von. 1867. *Cosmos: A Sketch of a Physical Description of the Universe*. New York: Harper & brothers.

James FC, McCulloch CE (1990) Multivariate analysis in ecology and systematics: panacea or Pandora's box? Ann Rev Ecol and Syst 21: 129-66

Jennings, S., Kaiser, M.J., & Reynolds, J.D. 2001.Marine fisheries ecology. Blackwell,Oxford

Jones, G.P., Planes, S. & Thorrold, S.R., 2005. Coral Reef Fish Larvae Settle Close to Home. *Current Biology*, 15(14), pp.1314-1318.

Kobayashi, D.R., 2006. Colonization of the Hawaiian Archipelago via Johnston Atoll: a characterization of oceanographic transport corridors for pelagic larvae using computer simulation. *Coral Reefs*, 25(3), pp.407-417.

Kosaki, Randall K, Richard L Pyle, John E Randall, and Darby K Irons. New Records of Fishes from Johnston Atoll, with Notes on Biogeography. Article. http://scholarspace.manoa.hawaii.edu/handle/10125/519.

Legendre P, and Legendre L. 1998. Numerical ecology. Developments in environmental modeling 20. Elsevier Science B.V., Amsterdam

Lester, Sarah E, and Benjamin I Ruttenberg. 2005. The relationship between pelagic larval duration and range size in tropical reef fishes: a synthetic analysis. *Proceedings of the Royal Society B: Biological Sciences* 272, no. 1563 (March 22): 585 -591.

Lester, Sarah E, Benjamin I Ruttenberg, Steven D Gaines, and Brian P Kinlan. 2007. The relationship between dispersal ability and geographic range size. *Ecology Letters* 10, no. 8 (August 1): 745-758.

Link JS, Brodziak JKT, Edwards SF, Overholtz WJ, Mountain D, Jossi JW, Smith TD, Fogarty MJ (2002) Marine ecosystem assessment in a fisheries management context. Can Jour of Fish & Aqua Sci 59(9): 1429-1440

Lobel, P. S. 1997. Comparative settlement age of damselfish larvae (Plectroglyphidodon imparipennis, Pomacentridae) from Hawaii and Johnston Atoll. *The Biological Bulletin* 193, no. 2: 281.

Lobel PS, Lobel LK (2004) Annotated checklist of the fishes of Wake Atoll. Pac Sci 58 (1): 65-90

MacArthur, Robert H., and Edward O. Wilson. 1967. *The Theory of Island Biogeography*. Princeton University Press.

Malcolm, H.A. Foulsham, E., Pressey, R.L. Jordan, A., Davies, P.L., Ingleton, T. Johnstone, N., Hessey, S.2012. Selecting zones in a marine park: Early systematic planning improves cost-efficiency; combining habitat and biotic data improves effectiveness. *Ocean & Coastal Management*, 59(0), pp.1–12.

Maragos, JE, Potts DC, Aeby GA, Gulko D, Kenyon J, Siciliano D, VanRavenswaay D. 2004. 2000–2002 Rapid ecological assessment of corals (Anthozoa) on shallow reefs of the Northwestern Hawaiian Islands. Part 1: species and distribution. Pac Sci 58: 211–230.

Mayr, Ernst. 1942 *Systematics and the Origin of Species*, Columbia Press: New York.

Mayr, Ernst. 1969. *Animal Species and Evolution*. Belknap Press of Harvard University Press.

Meekan MG, Choat JH (1997) Latitudinal variation in abundance of herbivorous fishes: a comparison of temperate and tropical reefs. Mar Biol 128: 373-383

Mora,C., P. M. Chittaro, P. F. Sale, J. P. Kritzer, and S.A. Ludsin. 2003. Patterns and processes in reef fish diversity. *Nature* 421,933–936.

Myers, R.F., 1999. *Micronesian Reef Fishes: A Field Guide for Divers and Aquarists* 1st ed., Coral Graphics.

NOAA (2006) Revision of species of concern list. Federal Register: 71 (200): 61022

R Development Core Team (2004) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3, URL HTTP://www.R-project.org.

Randall, J.E., 1996. *Fishes of the Great Barrier Reef & Coral Sea* Revised., Univ of Hawaii Pr.

Randall, J.E., 2002. *Surgeonfishes of the World*, Mutual Publishing.

Randall, J.E., 2005. *Reef and Shore Fishes of the South Pacific: New Caledonia to Tahiti and the Pitcairn Islands*, University of Hawaii Press.

Randall, J.E., 2007. *Reef and Shore Fishes of the Hawaiian Islands* 1st ed., University of Hawaii Sea Grant College Program.

Rivera MAJ, Kelley CD, Roderick GK. 2004. Subtle population genetic structure in the Hawaiian grouper, *Epinephelus quernus* (Serranidae) as revealed by mitochondrial DNA analyses. *Biological Journal of the Linnaean Society*. 81: 449–468.

Rooker, J. R., and G. D. Dennis. 1991. Diel, lunar and seasonal changes in a mangrove fish assemblage off southwestern Puerto Rico. *Bulletin of Marine Science* 49, no. 3: 684–698.

Ross, R.M., 1978. Territorial Behavior and Ecology of the Anemonefish Amphiprion melanopus on Guam1). *Zeitschrift für Tierpsychologie*, 46(1), pp.71-83.

Ross, R.M., 1978b. Reproductive behavior of the anemonefish Amphiprion melanopus on Guam. *Copeia*, pp.103–107.

Sale, P.F., 1971. Extremely Limited Home Range in a Coral Reef Fish, Dascyllus aruanus (Pisces; Pomacentridae). *Copeia*, 1971(2), pp.324-327.

Sale, P.F., 2002. *Coral reef fishes: dynamics and diversity in a complex ecosystem*, Academic Press.

Sale, P. F., P. J. Doherty, G. J. Eckert, W. A. Douglas, and D. J. Ferrell. 1984. Large scale spatial and temporal variation in recruitment to fish populations on coral reefs. *Oecologia* 64, no. 2 (October): 191-198.

Sale, Peter F., and William A. Douglas. 1984. Temporal Variability in the Community Structure of Fish on Coral Patch Reefs and the Relation of Community Structure to Reef Structure. *Ecology* 65, no. 2 (April 1): 409-422.

Sammarco, P.W. and Andrews, J.C., 1988. Localized Dispersal and Recruitment in Great Barrier Reef Corals: The Helix Experiment. *Science*, 239(4846), pp.1422–1424.

Sano, M., Shimizu, M. & Nose, Y., 1987. Long-term effects of destruction of hermatypic corals by Acanthaster planci infestation on reef fish communities at Iriomote Island, Japan. *Mar. Ecol. Prog. Ser*, 37(2/3), pp.191–199.

SAS Institute Inc., SAS 9.1.3 Help and Documentation, Cary, NC: SAS Institute Inc., 2000-2004.

Schultz, J.K. et al., 2006. Genetic connectivity among color morphs and Pacific archipelagos for the flame angelfish, Centropyge loriculus. *Marine Biology*, 151, pp.167-175.

Shanks, A.L., 2009. Pelagic Larval Duration and Dispersal Distance Revisited. *The Biological Bulletin*, 216(3), pp.373–385.

Simberloff, Daniel S. 1974. Equilibrium Theory of Island Biogeography and Ecology. *Annual Review of Ecology and Systematics* 5 (January 1): 161-182.

Simberloff, Daniel S., and Lawrence G. Abele. 1976. Island Biogeography Theory and Conservation Practice. *Science* 191, no. 4224 (January 23): 285 -286. doi:10.1126/science.191.4224.285.

Simberloff, Daniel S., and Edward O. Wilson. 1969. Experimental Zoogeography of Islands: The Colonization of Empty Islands. *Ecology* 50, no. 2 (March 1): 278-296.

doi:10.2307/1934856.

———. 1970. Experimental Zoogeography of Islands. A Two-Year Record of Colonization. *Ecology* 51, no. 5: 934-937. doi:10.2307/1933995.

Smith, J.M., 1966. Sympatric Speciation. *The American Naturalist*, 100(916), pp.637–50.

Syms, Craig, and Geoffrey P. Jones. 2000. Disturbance, habitat structure, and the dynamics of a coral reef fish community. *Ecology* 81, no. 10 (October): 2714-2729.

Thresher, R. E. and Brothers, E. B. 1985. Reproductive ecology and biogeography of Indo-West Pacific angelfishes (Pisces: Pomacanthidae). *Evolution* 39: 878–887.

Thresher, R. E., P. L. Colin, & L. J. Bell. 1989. Planktonic duration, distribution and population structure of western and central Pacific damselfishes (Pomacentridae). *Copeia* 1989: 420–434.

Via, Sara. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution* 16, no. 7 (July 1): 381-390. doi:10.1016/S0169-5347(01)02188-7.

Victor, B. C, and G. M Wellington. 2000. Endemism and the pelagic larval duration of reef fishes in the eastern Pacific Ocean. *Marine Ecology Progress Series* 205: 241–248.

Wallace, Alfred Russel. 1906. *The Malay Archipelago: The Land of the Orang-Utan and the Bird of Paradise; a Narrative of Travel, with Studies of Man and Nature*. London: Macmillan.

Ward, T.J., Vanderklift, M. A., Nicholls, A. O., and Kenchington, R. A. 1999. Selecting marine reserves using habitats and species assemblages as surrogates for biological diversity. *Ecological Applications*, 9(2), pp.691–698.

Weersing, K., and R. J Toonen. 2009. Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series* 393: 1–12.

Wellington, G. M. and B. C. Victor. 1989. Planktonic larval duration of one hundred species of Pacific and Atlantic damselfishes (Pomacentridae). *Marine Biology*. 101,557–568.

Zapata, F. A, and P. A. Herrón. 2002. Pelagic larval duration and geographic distribution of tropical eastern Pacific snappers (Pisces: Lutjanidae). *Marine Ecology Progress Series* 230: 295–300