January 2010

Phylogenetics and Phylogeography of Moray Eels (Muraenidae)

Joshua Reece
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Program in Evolution, Ecology and Population Biology

Dissertation Committee:
Allan Larson, Chairperson
Garland E. Allen
Jonathan M. Chase
Kenneth M. Olsen
Barbara Schaal
Alan R. Templeton

PHYLOGENETICS AND PHYLOGEOGRAPHY OF MORAY EELS

(MURAENIDAE)

by

Joshua Steven Reece

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment for the degree
of Doctor of Philosophy

August 2010

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Phylogenetics and Phylogeography of Moray Eels (Muraenidae)

by

Joshua S. Reece

Doctor of Philosophy

Division of Biology and Biomedical Sciences

Program in Evolution, Ecology, and Population Biology

Washington University, St. Louis 2010

This dissertation describes the evolutionary history of Moray eels (Muraenidae). Moray eels are apex predators on coral reefs around the world, but they are not well studied because of their cryptic habitats and occasionally aggressive behaviors. I provide the first molecular phylogeny of moray eels with widespread taxonomic and geographic coverage, including 44 species representing two subfamilies, eight genera, and all tropical ocean basins. Phylogenetic relationships among these taxa are estimated from portions of mitochondrial loci *cytochrome b* and *cytochrome oxidase subunit 1*, and portions of the nuclear loci *RAG-1* and *RAG-2*. I test four sets of contrasting phylogenetic hypotheses using standard topological tests. First, my phylogeny statistically supports the taxonomic distinction between true morays (Muraeninae) and snakemorays (Uropterygiinae). Second, I demonstrate that the durophagous characters (shell crushing jaws) of the genera *Gymnomuraena* and *Echidna* are not homologous. Third, I demonstrate that durophagous

ii

morphological characters have evolved in parallel in an ancestor of *Gymnomuraena* and at least three additional times within the genus *Echidna*. Finally, the tree topology indicates multiple invasions of the Atlantic from the Indo-Pacific, one of these occurring immediately prior to formation of the Isthmus of Panama approximately 2.8 MY ago and one or two others occurring in the early to mid Miocene. Cladogenesis occurring within the Atlantic during the mid Miocene and Pliocene also contributed to moray species diversity. These data include a pair of sister species separated by the Isthmus of Panama, allowing a time-calibrated tree with an estimated crown age for Muraenidae at between 41 and 60 MY ago, consistent with fossil evidence. Most lineage accumulation within morays occurred from the late Pliocene (~25 MY ago) through the Miocene (5-23 MY ago).

This dissertation also examines phylogeographic patterns of diversification within morays. Reef fishes disperse primarily as oceanic "pelagic" larvae, and debate continues over the extent of this dispersal, with recent evidence for geographically restricted (closed) populations in some species. In contrast, moray eels have the longest pelagic larval stages among reef fishes, possibly providing opportunities to disperse over great distances. I test this prediction by measuring mtDNA and nuclear DNA variation in two species of moray eels, *Gymnothorax undulatus* (N = 165) and *Gymnothorax flavimarginatus* (N = 124), sampled at 14-15 locations across the Indo-Pacific. The mtDNA data comprise 632 bp of *cytochrome b* and 596 bp of *cytochrome oxidase I.* Nuclear markers include two recombination-activating loci (421 bp of *RAG-1* and 754 bp of *RAG-2*). Analyses of molecular variance (AMOVA) and Mantel tests indicate little or no genetic differentiation, and no isolation by distance, across 22,000 km of the Indo-

Pacific. I estimate that mitochondrial genetic variation coalesces within the past ~2.3 million years for *G. flavimarginatus* and within the past ~5.9 million years for *G. undulatus*. Permutation tests of geographic distance on the mitochondrial haplotype networks indicate recent range expansions for some younger haplotypes (estimated within ~600,000 years) and episodic fragmentation of populations at times of low sea level. My results support the predictions that the extended larval durations of moray eels enable ocean-wide genetic continuity of populations. This is the first phylogeographic survey of the moray eels, and morays are the first reef fishes known to be genetically homogeneous across the entire Indo-Pacific.

Finally, this dissertation uses comparative phylogeography to examine the effects of niche breadth on population genetic structure. The effects of niche breadth on range size, rarity, and extinction risk have been well explored. However, the ability of niche breadth to affect population structure within species has never been examined in a comparative framework, possibly due to the unique set of conditions necessary for such a test. Population structure is often a result of differential gene flow among populations, and gene flow can be affected by both the vagility of an organism or its gametes (dispersal capability), or by the variety of habitats that an organism can occupy within its dispersal capabilities (niche breadth). While niche breadth has been shown to affect range size and rarity more than dispersal capability, it is unclear if this pattern extends to population genetic structure within species. I make use of a unique characteristic of moray eel (Muraenidae) biogeography. Moray eels are cosmopolitan species with the lowest level of endemism among the 4000+ species of reef fishes, and while many species exist in a variety of habitats throughout their range, other species with identical

ranges occupy only a small subset of those habitats, and have a diet that is also a specialized subset of the habitat generalists' diet. I compare measures of population genetic structure in two broad-niche species of moray eels (*Gymnothorax flavimarginatus, G. undulatus*) with two co-distributed but more specialized narrow-niche species of moray eels (*Echidna nebulosa* and *Gymnomuraena zebra)* throughout the Indo-Pacific. I report molecular genetic data for *E. nebulosa* and *G. zebra* using molecular markers orthologous to those used for *G. flavimarginatus* and *G. undulatus* for geographically overlapping sampling localities. I estimate geographic distributions of the four moray species using geographic information systems to delimit the distributions of habitats favorable for each species. Despite an 80% reduction in available habitat within the narrow-niche species and comparable dispersal capabilities, each of the four species shows high levels of connectivity throughout their respective ranges. These results indicate that in broadly-distributed species with high dispersal capabilities, niche breadth may have little to no effect on population structure.

Acknowledgments

Finally, I would like to thank the staff and faculty at Washington University in Saint Louis for their incredible kindness, longsuffering, and unheard-of efficiency. I have absolutely no desire to ever work with another group of people. Special thanks to Gerry Rohde, Linda Hurt, Melissa Torres, and Judy Musick. Andrew Johnstone has also provided unending support; I do not know of another institution that supports its graduate students with such zeal.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

PAGES 1-31

INTRODUCTION

Much of evolutionary biology concerns the patterns and processes that create and maintain biodiversity. Phylogeography is the study of the geographic distribution of genetic variation within species to reconstruct the origins and geographic expansion of lineages (Avise, 2000). Phylogenetics is the reconstruction of relationships among lineages (Futuyma 1998). Studies that incorporate phylogenetic relationships among species with phylogeographic patterns within species offer a comprehensive assessment of the microevolutionary and macroevolutionary dimensions of biodiversity.

My objectives are to describe patterns of biodiversity within moray eels (Muraenidae). I diagnose species-level lineages and examine the evolutionary history and timing of diversification within this group. I identify geographic patterns of connectivity among isolated oceanic reef populations in four species of moray eels, and in doing so assess how niche breadth affects geographic-genetic associations. My findings have broad implications for our understanding of marine fish biodiversity in the Indo-Pacific by identifying levels of genetic connectivity among oceanic reef populations and associated spatial and temporal patterns in species diversity.

Moray Eels (Muraenidae) are taxonomically nested within the monophyletic order Anguilliformes. This order comprises all "true eels" that share the synapomorphy of a particular larval form called a leptocephalus. Anguilliform fishes are the sister taxon to remaining teleost fishes and are known in the fossil record from the Middle Cretaceous, approximately 80 million years ago (Nelson, 2006). Monophyly of this group is supported by molecular and morphological data (Inoue et al., 2001, 2003). The order contains 15 families and 738 species (Nelson, 2006). Moray eels fall within a single family (Muraenidae) and contribute approximately 200 of the 738 species in the order

Anguilliformes, and the second largest taxonomic family behind the deep-water opichthid eels (Opichthidae, 250 species) (Nelson, 2006). Moray eels are known from the fossil record in the Miocene, approximately 34-54 million years ago (Benton, 1993), and are characterized by a long, scale-less body, eel-like shape, a dorsal fin with an origin usually before the gill openings, median fins confluent with the caudal fin, and no pelvic or pectoral fins. Gill openings are small roundish lateral openings, and the head has one to three lateral-line pores that are absent from the rest of the body. Moray eels range from 0.3 to 3.0 meters in total adult body length. They occur worldwide in tropical and temperate seas and act as top predators on coral-reefs, consuming fish, crustaceans and cephalopods (Böhlke et al., 1989). Adults are highly philopatric and territorial, and their larvae are generally widespread and abundant in pelagic trawl surveys; adults range in depth from 1-200 meters below sea level (Böhlke et al., 1989). Moray eels also have the longest pelagic larval durations (a proxy for dispersal capability) (Lester and Ruttenberg, 2005) of any coral-reef fish (upwards of 80 days) and the lowest level of endemism among coral-reef fish taxonomic families (Hourigan & Reese 1987).

Three chapters of this dissertation present new data and analyses. Each chapter contains an introduction to the material discussed, separate figures and tables, and a separate literature-cited section. The first chapter examines the evolutionary history and timing of lineage accumulation in moray eels. The second chapter describes phylogeographic structure and the role of pelagic larval duration in the population genetics of two common moray eel species, *Gymnothorax undulatus* and *Gymnothorax flavimarginatus*. The third chapter examines population structure in two additional moray eel species, *Gymnomuraena zebra* and *Echidna nebulosa*, which differ from the previous

two *Gymnothorax* species by having a more specialized set of habitat tolerances and diet. In this chapter I test the hypothesis that reduced niche breadth creates phylogeographic population structure throughout the Indo-Pacific.

Moray eels are a poorly studied group of organisms with respect to their evolutionary history and patterns of gene flow, yet they are one of the most widespread and common groups of reef fishes. Much of this work builds upon decades of life history, census, and biogeographic studies of morays, and none of this work would have been possible without invaluable museum collections of specimens, taxonomic expertise, location records, and DNA vouchers.

Literature Cited

Avise, J.C., 2000. Phylogeography: The history and formation of species. Harvard University Press, Cambridge, MA.

Benton, M.J. (Ed.), 1993. The Fossil Record 2. Chapman & Hall, London.

Böhlke, E.B., McCosker, J.E., Böhlke, J.E., 1989. Family Muraenidae. In: Böhlke, E.B. (Ed.), Fishes of the Western North Atlantic. Sears Foundation for Marine Research, New Haven.

Futuyma, D.J. 1998. Evolutionary Biology. 3rd ed. Sinauer Associates, Sunderland, Massachusetts.

Hourigan, T.F. 1987. Mid-ocean isolation and the evolution of Hawaiian reef fishes. Trends in Ecology and Evolution 2, 187-191.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2001. A mitogenomic perspective on the basal teleostean phylogeny: resolving higher-level relationships with longer DNA sequences. Molecular Phylogenetics and Evolution 20, 275-285.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2003. Basal actinopterygian relationships: a mitogenomic perspective on the phylogeny of the "ancient fish". Molecular Phylogenetics and Evolution 26, 110-120.

Lester, S.E., Ruttenberg, B.I., 2005. The relationship between pelagic larval duration and range size in tropical reef fishes: a synthetic analysis. Proceedings of the Royal Society B 272, 585-591.

Nelson, G.J., 2006. Fishes of the World 4th Edition. John Wiley & Sons, Inc., Hoboken, New Jersey.

CHAPTER 2

PAGES 6-31

**MOLECULAR PHYLOGENETICS OF MORAY EELS (MURAENIDAE)
DEMONSTRATES MULTIPLE ORIGINS OF A SHELL-CRUSHING JAW
(*GYMNOMURAENA, ECHIDNA*) AND MULTIPLE COLONIZATIONS OF THE
ATLANTIC OCEAN**

Introduction

Moray eels (Muraenidae) are a taxonomic family of approximately 200 species of predatory reef fishes found in every tropical ocean basin. Muraenid taxonomy was first reviewed by Günther (1870), and more recent regional taxonomic reviews have been conducted for the Atlantic (Böhlke et al., 1989), Hawaii (Böhlke and Randall, 2000), the broader Indo-Pacific (Böhlke and Smith, 2002), the Red Sea (Randall and Golani, 1995), and Australia (Böhlke and McCosker, 2001). Jiménez *et al.* (2007) and Almada *et al.* (2009) evaluated evolutionary relationships among Mediterranean and northeastern Atlantic moray eel species, respectively; however, neither study evaluated generic or higher order taxonomic relationships within moray eels. Moray eels are currently divided into two subfamilies, Muraeninae (morays) and Uropterygiinae (snakemorays). The distinguishing morphological characteristics of Uropterygiinae include dorsal and anal fins restricted to the tip of the tail; in Muraeninae, the dorsal fin usually begins near the gill opening, and the anal fin begins just posterior to the anus, approximately mid-body (Böhlke et al., 1989). The subfamily Uropterygiinae contains four genera and 36 species (Loh et al., 2008), whereas Muraeninae contains nine genera and approximately 166 species (Böhlke et al., 1989). Nelson (1966) considers the deossification and loss of some hypobranchial elements from the pharyngeal jaws a shared derived character of Muraeninae that distinguishes this subfamily from Uropterygiinae and other eels (order Anguilliformes). Mitochondrial DNA data from west Pacific species support the phylogenetic distinction between Uropterygiinae and Muraeninae (Loh et al., 2008). Moray eel species often occur sympatrically and can be diagnosed with varying levels of

resolution by the positions of dorsal and anal fins, tooth placement and morphology, larval pigmentation, adult color pattern, and vertebral numbers (Böhlke et al., 1989).

Mehta (2008) described two morphotypes within Muraenidae corresponding to piscivorous (fish eating) and durophagous (shell-crushing) jaws. These morphotypes differ in their cranial and pharyngeal jaw structures and most noticeably in their dentition. Piscivorous species have elongate oral jaws with numerous long, recurved fang-like teeth. The oral jaws of durophagous morays are short, recurved, and covered with numerous molariform teeth for crushing shelled prey. Dietary studies confirm that species classified by their cranial morphology, oral jaws and dentition (as described above) as durophagous feed primarily on crustaceans, whereas those with piscivorous morphology feed primarily or exclusively on fishes (literature reviewed in Table 1 of Mehta 2008). Taxonomically, the piscivorous morphology characterizes the largest genus of Muraeninae (*Gymnothorax*) and most of the remaining genera. The durophagous morphology occurs in the genus *Echidna* (with 10 recognized species) and the monotypic genus *Gymnomuraena*. It is unknown whether the durophagous morphology is homologous among these two genera and 11 species, or whether they include parallel origins of a durophagous jaw from a (presumed) piscivorous ancestor.

Muraenidae includes approximately 150 Indo-Pacific species and 50 Atlantic species (Böhlke et al., 1989; Böhlke and Smith, 2002). Most reef fishes have a center of diversity in the Indo-Pacific (Briggs, 1999), which potentially includes the sites of origin of many taxa. Butterflyfishes (Fessler and Westneat, 2007) and wrasses of the family Labridae (Westneat and Alfaro, 2005) show multiple invasions from the Indo-Pacific into the Atlantic. Conversely, Atlantic blennies of the genus *Ophioblennius* (Muss et al.,

8

2001), pygmy angelfishes of the genus *Centropyge* (Bowen et al., 2006) and the wrasse genera *Halichoeres* (Barber and Bellwood, 2005) and *Thalassoma* (Bernardi et al., 2004), represent single colonization events from the Indo-Pacific via southern Africa. At least three well-surveyed groups of reef fishes (damselfish, some wrasses, and parrotfishes) show multiple invasions of and local diversification within the western Atlantic (Rocha et al., 2008). Multiple genera of moray eels occur in the Atlantic, and the widespread genus *Gymnothorax* occurs in every ocean basin. It is unclear whether the occurrence of *Gymnothorax* species in the Atlantic represents multiple invasions from the Indo-Pacific or a single invasion and subsequent speciation in the Atlantic.

I generate a phylogeny of 44 moray eel species to test four major hypotheses about moray eel evolutionary history and biogeography. First, I test hypotheses of monophyly for the morphologically diagnosable subfamilies Muraeninae and Uropterygiinae. Second, I test the hypothesis that the durophagous genera *Gymnomuraena* and *Echidna* form a monophyletic group and that their evolutionarily derived feeding morphologies are homologous. Prior taxonomic grouping of the monotypic genus *Gymnomuraena* with *Echidna* is based primarily on shared durophagy and not well supported by other morphological characters. Third, I test whether the genus *Echidna* is monophyletic and whether durophagy is homologous among species within the genus. Finally, I test the hypothesis that Atlantic members of *Gymnothorax* form a monophyletic group and represent a single invasion from the Indo-Pacific. Figure 2-1 summarizes expected topologies corresponding to each of the four hypotheses. I use fossil and biogeographic data to produce a time-calibrated phylogeny to test these hypotheses, including estimates of a crown age for the extant moray lineages, the

evolution of durophagy, and approximate timing of inter-oceanic exchanges between the Atlantic and Indo-Pacific.

Materials and methods

Sequence generation

I acquired specimens from tissue banks, personal collections, collaborators, and the pet trade when voucher and locality information were available. Dr. David G. Smith at the United States National Museum confirmed the identity of a voucher specimen for each DNA sequence in this study. I extracted DNA using Viogene DNA Kits (www.viogene.com) and manufacturer's protocols. Polymerase chain reactions featured a total volume of 25µL including 5µL of Promega (www.promega.com) 5x buffer, 2.5µL of 25mM $MgCl_2$, 2.5µL of 0.2µM dNTPs, 2.5µL of 0.2µM of each primer, 0.125µL (1 unit) of Promega GoTaq DNA polymerase, and 2µL of template DNA at approximately 5ng/µL. A 632-bp fragment of *cytochrome b* (*CYB*) was amplified using the primers L14725 (5'-GTG ACT TGA AAA ACC ACC GTT G-3') (Song et al., 1998) and H15573 (5'-AAT AGG AAG TAT CAT TCG GGT TTG ATG-3') (Taberlet et al., 1992) and an annealing temperature of 50°C. A 596-bp fragment of *cytochrome oxidase subunit 1* (*COI*) was amplified using primers FishF2 (5'-TCG ACT AAT CAT AAA GAT ATC GGC AC-3') and FishR2 (5'-ACT TCA GGG TGA CCG AAG AAT CAG AA-3') (Ward et al., 2005) and annealing temperature of 50°C. A 421-bp fragment of the nuclear recombination activation gene *RAG-1* was amplified using primers RAG1-F3 (5'-GCC TCA GAA AAC ATG GTG CT-3') and RAG1-R3 (5'-CCA CAC AGG TTT CAT CTG GA-3') (Reece et al. 2010; Chapter 3) with an annealing temperature of 50°C. A 754-bp

fragment of the nuclear recombination activation gene *RAG-2* was amplified using

primers RAG2-F3 (5'-AGG TGA CCC TTC GTT GTC AG-3') and RAG2-R3 (5'-ATG

AGG CTC CCT TCC AAA GT-3') (Reece et al. 2010; Chapter 3) at an annealing

temperature of 52°C. The thermal profiles for PCR were 95°C for 3 min, followed by 35

cycles of 95°C for 1 min, annealing temperature for 40 sec, and 72°C for 45 sec, with a

final elongation at 72°C for 7 min.

PCR products were visualized through 1.5% agarose gel electrophoresis and

purified using Exo-Sap or Viogene Gel Purification Kits using manufacturer's protocols.

Sequences were generated on ABI 3130 and ABI 3330 Automated DNA sequencers at

the Washington University Genome Sequencing Center and the Smithsonian Museum

Support Center using PCR primers listed above. DNA sequences were manually edited

using Sequencher v.4.8, and aligned by hand. I performed all sequencing and editing. For

nuclear markers, heterozygous positions were identified by a secondary peak in the

electropherograms reaching at least 25% of the intensity of the primary peak. Gametic

phases of nuclear sequences with more than a single heterozygous site were estimated

using a Bayesian approach implemented in the software program Phase v.2.1 (Stephens

and Donnelly, 2003; Stephens et al., 2001). All Phase analyses were run through five

iterations with different random-number seeds and run for 1000 iterations with a single

thinning interval and 100 burn-in iterations. Consistency of results was determined by

examining allele frequencies and coalescent goodness-of-fit measures estimated for each

of the four runs. If haplotypes could not be estimated with 90% posterior probability,

each ambiguous site was coded as missing data. Less than 3% of all nucleotide characters

were coded as missing data under this criterion.

Phylogenetic analyses

I calculated Tajima's D (Tajima, 1989) test for selection in DNAsp v.5.1.0 (Rozas et al., 2003) to demonstrate that the loci included in this study were consistent with neutral expectations. Phylogenetic trees were constructed in the programs MrBayes v.3.1 (Ronquist and Huelsenbeck, 2003) and BEAST v.1.5.4 (Drummond and Rambaut, 2007). I used the Conger Eel (*Conger conger*) as an Anguilliform outgroup for combined portions of the mitochondrial genes *CYB* and *COI*, and for the nuclear genes *RAG-1* and *RAG-2*. A model of evolution was computed for three partitions within each gene region, corresponding to 1st, 2nd and 3rd codon positions, using jModelTest v.0.1.1 (Posada, 2008). Each of three MrBayes runs, corresponding to combined mtDNA, *RAG-1*, and *RAG-2*, consisted of 3,000,000 iterations of four chains replicated in two independent runs with a sampling interval of 100 iterations and burn-in of 7,500 runs. Each run reached stationarity with these values; ESS values were over 200 for each parameter estimated, and the standard deviations between independent runs had stabilized and were below 0.001. Stationarity of all runs was interpreted using Tracer v.1.5.0 (Rambaut and Drummond, 2007). To demonstrate that a concatenated dataset was appropriate for these gene regions, I conducted pairwise Shimodaira-Hasegawa (S-H) Tests (Shimodaira and Hasegawa, 1999) among the gene tree topologies (mtDNA, *RAG-1*, *RAG-2*). After all gene topologies were identified as congruent by S-H Tests, I used a concatenated dataset (maintaining the partitioning strategy described above for a total of 4 genes x 3 partitions within each gene = 12 partitions) to reconstruct evolutionary relationships using the same run parameters as above but for 10,000,000 generations. The phylogenetic reconstruction

executed in BEAST differed only in the use of time calibrations, and those methods are discussed below.

I tested four phylogenetic hypotheses that correspond to 1) the two subfamilies of Muraeninae and Uropterygiinae are monophyletic, 2) the durophagous genera *Echidna* and *Gymnomuraena* form a monophyletic group with respect to the genus *Gymnothorax*, 3) genus *Echidna* is monophyletic, and 4) Atlantic species of *Gymnothorax* form a monophyletic group. The taxonomic sampling scheme used to address each of these tests is described below. Each of these hypotheses was tested by three metrics (in order of increasing stringency) to determine whether the data statistically discriminate hypotheses 1-4 from contrasting alternatives: Bayes Factors (Kass and Raftery, 1995), S-H Tests (Shimodaira and Hasegawa, 1999), and Templeton tests (Templeton, 1983). Bayes Factors were calculated by comparing the logarithms of overall likelihood scores of the favored topology to the best topology compatible with the contrasting hypothesis. This value was then doubled and interpreted as the Bayes Factor ($2logB_{10}$) according to Table 2 in Kass and Raftery (1995). A Bayes Factor exceeding 10 is considered definitive rejection of the less likely topology (Kass and Raftery, 1995). S-H and Templeton tests were executed in PAUP* v.4.0 (Swofford, 2003). S-H tests with 1000 bootstrap replicates checked statistical significance of differences in likelihood values of the favored tree and those of the 100 most likely trees sampled from the Bayesian posterior probability distribution of the contrasting hypothesis. I considered a hypothesis rejected if all of the 100 best trees compatible with it were statistically rejected as less compatible with the data than the favored tree. For Templeton tests, I constrained a parsimony analysis to find the shortest tree compatible with hypotheses 1-4 or their converse as appropriate, and

then evaluated statistical significance of the differences in length of the contrasting pairs of topologies.

All species sampled in this study were used to test each hypothesis, and the pertinent taxonomic coverage is as follows: To test for monophyly of Muraeninae and Uropterygiinae, I used samples from three species of two genera in Uropterygiinae: *Uropterygius macrocephalus*, *U. fuscoguttatus*, and *Scuticaria tigrina* and 41 Muraeninae species representing the genera *Gymnomuraena*, *Gymnothorax*, *Echidna*, *Enchelynassa*, *Rhinomuraena*, and *Enchelycore*. I tested monophyly of each subfamily and evaluated the possibility that either family is monophyletic and nested within a paraphyletic group comprising species from the other subfamily. My second hypothesis was that the durophagous species *Gymnomuraena zebra*, *Echidna nebulosa*, *E. polyzona*, *E. rhodochilus*, and *E. leucotaenia* form a monophyletic group relative to the 36 piscivorous species sampled. Monophyly of *Echidna* was tested using the four *Echidna* species relative to all other samples. Lastly, members of Indo-Pacific Muraeninae were compared to five Atlantic *Gymnothorax* species (*G. saxicola*, *G. vicinus*, *G. ocellatus*, *G. moringa*, and *G. miliaris*) to test the hypothesis of a single invasion into the Atlantic from the Indo-Pacific.

The program BEAST was used to estimate both a phylogenetic topology and the timing of cladogenic events based on two independent and cross-validated calibrations. The BEAST analysis included the same partitioning strategy and models of evolution as did the MrBayes run. A joint tree was estimated from the four gene trees, and molecular clock and mutation models were unlinked across all 12 partitions. A relaxed uncorrelated lognormal clock prior was used, along with a Yule speciation process, per

recommendations for interspecific phylogenies (Drummond and Rambaut, 2007). A uniform prior ranging between 2.8 and 3.1 MYA (million years ago) was used as a biogeographic calibration point for the TMRCA (time to most recent common ancestor) of *G. meleagris* and *G. miliaris*. These are suspected sister taxa based on larval and adult similarities, and are separated by the Isthmus of Panama, which created a land bridge between the Indo-Pacific and Atlantic most recently at approximately 2.8 to 3.1 MY (Bermingham et al., 1997; Duque-Caro, 1990; Lessios, 2008; Marko, 2002). A second calibration point corresponding to a crown age was translated into a flat prior between 50 and 65 MYA. This calibration is widely regarded as encompassing the diversification period of most major orders of tropical marine fishes, including eels, based on fossil evidence (Bellwood and Wainwright, 2002), and is consistent with fossil calibrations in several time-calibrated phylogenetic studies of marine fishes (Bellwood et al., 2004; Westneat and Alfaro, 2005). To cross-validate the calibrations, two additional runs were made with only one calibration as a prior, and the other estimated by the program. In both cases the proposed calibrations fell within the 95% confidence intervals of estimated values (see Results) and were used in the final analysis. As a final validation, I calculated the mutation rates for each of the *CYB* and *COI* genes using an estimated closing of the Panamanian Isthmus at 3.1 MY, the portion of the estimate range (2.8 to 3.1 MYA) historically used as a calibration point in other studies of marine fishes. Studies published prior to Lessios (2008) used a 3.1 to 3.5 MYA estimate for the most recent closing of the Panamanian Isthmus. The estimated mutation rates for *CYB* and *COI* are concordant with those reported for other marine fishes (see Results and Discussion), and are interpreted as being reasonable estimates for divergence times. Because nuclear gene divergences

(*RAG-1* and *RAG-2*) were also available for these sister taxa, and the 3.1 MY divergence

time was concordant with previous work (Bermingham et al., 1997), I calibrated a

species-specific mutation rate for the nuclear markers *RAG-1* and *RAG-2*. After trial

BEAST runs and corresponding modifications to the priors, two final runs were

completed with 1 billion iterations, sampling every 5000 steps. I joined the two runs

using LogCombiner 1.5.2 (Drummond and Rambaut, 2007) and produced a time-

calibrated phylogeny using TreeAnnotator 1.5.2 (Drummond and Rambaut, 2007).

Results

A total of 2403 bp of DNA was resolved, corresponding to 632 bp of the

mitochondrial locus *CYB* and 596 bp of *COI*; 421 bp of the nuclear gene *RAG-1* and 754

bp of the nuclear gene *RAG-2*. Tests for selection on mtDNA, *RAG-1*, and *RAG-2*

indicated neutrality for all loci. Corresponding Tajima's D values (*COI* = 0.075, *CYB* =

0.11, *RAG-1* = -1.3, *RAG-2* = -1.7) were non-significant at $P > 0.1$. All pairwise gene-tree

topological S-H tests were non-significant and support the use of concatenated, mixed-

model analyses. For each of the four hypotheses described in Figure 2-1, all three tests

(Bayes Factors, S-H, and Templeton tests) were concordant. Results in Figure 2-1

indicate whether the data are compatible with the stated hypothesis or reject it in favor of

a contrasting topology (Figure 2-2).

My phylogenetic tests support my first hypothesis of monophyly of subfamilies

Muraeninae and Uropterygiinae. The most parsimonious tree constrained to keep

Uropterygiinae non-monophyletic has 3211 steps, and the most parsimonious tree for a

non-monophlytic Muraeninae has 3226 steps, both significantly longer than the favored

tree (3115 steps) using S-H and Templeton tests ($P < 0.01$ for each comparison) and

Bayes Factors (BF=84, 98, respectively). Within Muraeninae, the durophagous *Gymnomuraena zebra* (monotypic genus) is the sister taxon to a clade comprising all other muraenids. My second hypothesis that *Gymnomuraena* and *Echidna* form a monophyletic group is rejected with all three tests of topological concordance (Bayes Factors = 64; $P < 0.01$ for S-H and Templeton tests). Of the four *Echidna* species sampled, *E. leucotaenia* and *E. polyzona* form a clade, but *E. nebulosa* and *E. rhodochilus* are phylogenetically distantly removed from these species and from each other (Figure 2-2). My third hypothesis of monophyly of *Echidna* is statistically rejected (Bayes Factors = 28; $P < 0.01$ for S-H and Templeton tests). The durophagous condition appears to have evolved at least four times in muraenid evolutionary history.

The five species of Atlantic *Gymnothorax* include two species pairs but do not form a monophyletic group with respect to Indo-Pacific species. Topological tests uniformly reject my fourth hypothesis of monophyly of Atlantic *Gymnothorax* species (Bayes Factors = 96; $P < 0.01$ for S-H and Templeton tests). My current sampling indicates a minimum of three *Gymnothorax* invasions from the Indo-Pacific into the Atlantic, and the actual number is likely to be much higher.

The time-calibrated phylogeny produced in BEAST (Figure 2-2) is based on two calibration points, cross-validated in three ways: 1) phylogenies created using only the 2.8 to 3.1 MY calibration estimated a TMRCA for all Muraenidae at between 33 and 74 MY, fully encompassing the external calibration of 50 to 65 MY. 2) Phylogenies using only the 50-65 MY calibration for all Muraenidae estimated a TMRCA between *G. miliaris* and *G. meleagris* at 1.4 to 5.2 MY, which encompasses the 2.8 to 3.1 MY calibration. 3) Using a divergence time of 3.1 MY (Duque-Caro, 1990; Lessios, 2008) for

the sister taxa *G. miliaris* and *G. meleagris*, I calculate a mutation rate of 2.1%

divergence per MY (between lineages) for portions of the *CYB* gene and 1.2% for

portions of the *COI* gene; both values mirror those reported in the literature for marine

fishes (see Discussion). Calibrations for *RAG-1* and *RAG-2* based on combined mtDNA

divergences give estimated rates of 0.45% and 0.35% divergence per MY (between

lineages), respectively.

Discussion

Despite the potential difficulties in identifying and classifying moray species

(Randall, 2007), my results support the deepest morphology-based taxonomy within

Muraenidae; subfamily classifications of Muraeninae and Uropterygiinae. These

subfamilies are distinguished by a suite of diagnostic characteristics, including restriction

of the anal and dorsal fins to the posterior tip of the tail in Uropterygiinae and their

extension beyond this region in Muraeninae. Based on my current sampling of

Muraenidae, the two subfamilies are monophyletic and sister taxa. A previous study with

more thorough sampling of Uropterygiinae (Loh et al., 2008) is compatible with this

conclusion.

The zebra moray, *Gymnomuraena zebra*, represents a monotypic genus

characterized by a strongly ossified skeleton, enlarged dermal bones, and posterior

placement of the anus. *Gymnomuraena zebra* and members of the genus *Echidna* feed

almost exclusively on crustaceans, and all share common morphological features

associated with this diet, including rounded, molariform teeth, and short, recurved jaws.

Gymnomuraena zebra appears to have diverged from a common ancestor of all other

Muraeninae approximately 43 MY ago, and based on my phylogenetic inferences, its

durophagous feeding and associated morphological characters evolved separately from those of other durophagous muraenines. Nonmonophyly of the durophagous genus *Echidna* indicates that durophagous feeding and associated jaw morphology evolved at least three additional times in Muraeninae; six unsampled species of *Echidna*, including one in the Atlantic, might include additional origins of durophagous feeding. Morphology-based taxonomy in this case is confounded by parallel evolution of durophagy and associated jaw morphology in Muraeninae. Remarkably, the nominal *Echidna* species in my survey are separated by about 20 MY. This finding will warrant further morphological investigations and potential revision of the genus *Echidna* to reflect a monophyletic taxonomic grouping.

Taxonomic families and subfamilies of reef fishes typically have broad or even cosmopolitan distributions with a center of species diversity in an area surrounding Indonesia, New Guinea and the Philippines, often called the Coral Triangle (Allen, 2008; Briggs, 2009; Veron et al., 2009). Species diversity decreases with distance from the Coral Triangle, and is almost universally lower in the Atlantic (e.g., Briggs, 1995; Myers, 1991; Springer, 1982; Veron, 1995). Some lineages of reef fishes have invaded the Atlantic multiple times (e.g., Bowen et al., 2006; Fessler and Westneat, 2007; Rocha et al., 2005), and others have diversified there following a single invasion (Muss et al., 2001). These invasions could occur from the proto-eastern Pacific prior to closure of the Isthmus of Panama 2.8 to 3.1 MYA (Duque-Caro, 1990; Lessios, 2008), from the Indian Ocean via southern Africa (Gordon, 2003), or through the now closed (15-20 MYA) Tethys Sea (Smith et al., 2004). The second pathway effectively closed to tropical fauna about 2.5 MYA with the onset of cold-water (Benguela) upwelling, but opened

intermittently at the end of each Pleistocene glacial cycle (Peeters et al., 2004; Shannon, 1985). The five Atlantic *Gymnothorax* species sampled represent at least two and probably three invasions from the Indo-Pacific. Cladogenesis associated with formation of the Isthmus of Panama explains the separation of the Atlantic *G. miliaris* from its Pacific sister species *G. meleagris* 2.9 MY ago. Two other pairs of Atlantic sister species are *G. moringa* and *G. vicinus* (14.2 MY; Figure 2-2) and *G. ocellatus* and *G. saxicola* (3.2 MY; Figure 2-2). My phylogenetic analysis indicates that the ancestral lineages of these two pairs of species each separated from its closest Indo-Pacific relatives in the early to mid Miocene (16-21 MY ago), coincident with the closing of the Tethys Sea. Although my favored topology indicates two separate invasions of the Atlantic in the early to mid Miocene, branch support is not sufficient to reject the alternative hypothesis of a single Atlantic invasion followed by return of a descendant lineage to the Indo-Pacific. The Atlantic contains most members of the unsampled genus *Muraena*, which is hypothesized to have been distributed across the Tethys Sea 15-20 MYA (Almada et al., 2009; Smith et al., 2004). This time frame is compatible with my inferences of the oldest invasions of the Atlantic by *Gymnothorax*. I conclude that recurring invasions from the Indo-Pacific and *in situ* speciation both contribute to Atlantic moray diversity (Figure 2-2).

Although my taxonomic sampling is too limited to estimate lineage-accumulation rates for moray eels, my results indicate that lineage accumulation was greatest in the Miocene (approximately 5-23 MY ago), with relatively few lineages predating a late Oligocene date of ~25 MY ago and few originations occurring in the Pliocene (approximately 2.5-5 MY ago). These dates are consistent with limited fossil evidence for

the first appearance of moray eels at 34 to 54 MYA (Benton, 1993), and of modern day *Gymnothorax*-type species in the Mediterranean as recently as 5.3 MYA (Arambourg, 1927; Gaudant, 2002). My work in Chapter 3 shows geographic genetic continuity within species of *Gymnothorax* (Reece et al., 2010), indicating that widespread discovery of cryptic species in moray eels is unlikely, in contrast to results obtained for many terrestrial and freshwater vertebrate taxa. My estimate that extant moray species trace their cladogenesis primarily to events occurring between the late Oligocene and Pliocene is unlikely to be biased by overlooking large numbers of younger, cryptic species, and is consistent with a broad taxonomic review of marine fish diversification (Rocha and Bowen, 2008). Results in Chapter 3 also indicate that some *Gymnothorax* species have likely maintained geographically widespread distributions throughout the Indo-Pacific for much of their evolutionary history. The estimated mitochondrial-haplotype coalescence time within *G. undulatus* is late Miocene (~5.9 MY ago), and phylogenetically disparate haplotypes are widely shared among the geographic populations of this species (Reece et al. 2010; Chapter 3). Moray eels appear to have accumulated high species diversity and worldwide distributions through a long history of lineage accumulation, especially within the last 25 million years.

Figure Legends

Figure 2-1. Expected topologies based on four null hypotheses. Topologies are presented in order of increasing resolution, with the groupings whose monophyly is being tested shown in bold type. All three metrics (Bayes Factors, S-H, and Templeton tests) indicate statistically significant rejection of hypotheses 2-4 and of the converse of hypothesis 1 using a concatenation of all molecular markers.

Figure 2-2. Bayesian phylogenetic reconstruction of the time-calibrated phylogeny of Muraenidae based on a concatenated dataset of portions of the *COI*, *CYB*, *RAG-1*, and *RAG-2* genes. The two subfamilies are labeled, and an asterisk at a node indicates Bayesian posterior probability support above 0.95. Divergence times are in millions of years, with gray bars denoting the 95% posterior probability densities around point estimates. Species of the genus *Gymnothorax* that are from the Atlantic are followed by "ATL" in bold type, and the species names for all durophagous species are in bold type.

Accept

Outgroup

Uropterygiinae

Muraeninae

Hypothesis 1: Subfamily taxonomy reflects evolutionary history

Reject

Outgroup

Uropterygiinae

Durophagous Muraeninae

Piscivorous Muraeninae

Hypothesis 2: Single Evolution of Durophagy, monophyly of *Gymnomuraena* and *Echidna*

Reject

Outgroup

Uropterygiinae

Durophagous Muraeninae (*Gymnomuraena*)

Durophagous Muraeninae (*Echidna*)

Piscivorous Muraeninae

Hypothesis 3: Monophyly of durophagous *Echidna*

Reject

Outgroup

Uropterygiinae

Durophagous Muraeninae (*Gymnomuraena*)

Durophagous Muraeninae (*Echidna*)

Atlantic Piscivorous *Gymnothorax*

Indo-Pacific Piscivorous Gymnothorax

Hypothesis 4: Single invasion into Atlantic, monophyly of Atlantic *Gymnothorax*

Figure 2-1

Figure 2-2

Literature Cited

Allen, G.R., 2008. Conservation hotspots of diversity and endemism for Indo-Pacific. Aquatic Conservation: Marine and Freshwater Ecosystems 18, 541-556.

Almada, V.C., Pérez-Portela, R., Robalo, J.I., Brito, A., 2009. Phylogenetic relationships of *Gymnothorax bacalladoi* (Böhlke and Brito 1987) a poorly known moray eel of the Macaronesian islands. Molecular Phylogenetics and Evolution 52, 252-256.

Arambourg, C., 1927. Les poissons fossiles d'Oran. Matériaux pour la carte géologique d'Algérie, 1ère série, Paléontologie 6, 1-298.

Barber, P.H., Bellwood, D.R., 2005. Biodiversity hotspots: evolutionary origins of biodiversity in the wrasses (*Halichoeres*: Labridae) in the Indo-Pacific and new world tropics. Molecular Phylogenetics and Evolution 35, 235-253.

Bellwood, D.R., van Herwerden, L., Konow, N., 2004. Evolution and biogeography of marine angelfishes (Pomocanthidae). Molecular Phylogenetics and Evolution 33, 140-155.

Bellwood, D.R., Wainwright, P.C., 2002. The history and biogeography of fishes on coral reefs. In: Sale, P.F. (Ed.), Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem. Academic Press, San Diego, pp. 5-32.

Benton, M.J. (Ed.), 1993. The Fossil Record 2. Chapman & Hall, London.

Bermingham, E., McCafferty, S.S., Martin, A.P., 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian isthmus. In: Kocher, T.D., Stepien, C.A. (Eds.), Molecular systematics of fishes. Academic Press, New York, pp. 113-128.

Bernardi, G., Bucciarelli, G., Costagliola, D., Robertson, D.R., Heiser, J.B., 2004. Evolution of coral reef fish *Thalassoma* spp. (Labridae). 1. Molecular phylogeny and biogeography. Marine Biology 144, 369-375.

Böhlke, E.B., McCosker, J.E., 2001. The moray eels of Australia and New Zealand, with the description of two new species (Anguilliformes: Muraenidae). Records of the Australian Museum 53, 71-102.

Böhlke, E.B., McCosker, J.E., Böhlke, J.E., 1989. Family Muraenidae. In: Böhlke, E.B. (Ed.), Fishes of the Western North Atlantic. Sears Foundation for Marine Research, New Haven.

Böhlke, E.B., Randall, J.E., 2000. A review of the moray eels (Anguilliformes: Muraenidae) of the Hawaiian Islands, with descriptions of two new species. Proceedings of the Academy of Natural Sciences of Philadelphia 150, 203-278.

Böhlke, E.B., Smith, D.G., 2002. Type Catalogue of the Indo-Pacific Muraenidae. Proceedings of the Academy of Natural Sciences of Philadelphia 152, 89-172.

Bowen, B.W., Muss, A., Rocha, L.A., Grant, W.S., 2006. Shallow mtDNA coalescence in Atlantic Pygmy Angelfishes (Genus *Centropyge*) indicates a recent invasion from the Indian Ocean. Journal of Heredity 97, 1-12.

Briggs, J.C., 1995. Global Biogeography. Elsevier, Amsterdam.

Briggs, J.C., 1999. Coincident biogeographic patterns: Indo-West Pacific Ocean. Evolution 53, 326-335.

Briggs, J.C., 2009. Diversity, endemism, and evolution in the Coral Triangle. Journal of Biogeography 36, 2008-2010.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by
sampling trees. BMC Evolutionary Biology 7, 214.

Duque-Caro, H., 1990. Neogene stratigraphy, paleoceanography and paleobiology in
northwest South America and the evolution of the Panama Seaway.
Palaeogeography, Palaeoclimatology, Palaeoecology 77, 203-234.

Fessler, J.L., Westneat, M.W., 2007. Molecular phylogenetics of the butterflyfishes
(Chaetodontidae): taxonomy and biogeography of a global coral reef fish family.
Molecular Phylogenetics and Evolution 45, 50-68.

Gaudant, J., 2002. La crise messinienne et ses effets sur l'ichthyofaune néogène de la
Méditerranée: le témoignage des squelettes en connexion de poissons téléostéens.
Geodiversitas 24, 691-710.

Gordon, A.L., 2003. The brawniest retroflection. Nature 421, 904-905.

Günther, A., 1870. Catalogue of the fishes in the British Museum. Catalogue of the
Physostomi, containing the families Gymnotidae, Symbranchidae, Muraenidae,
Pegasidae, and of the Lophobranchii, Plectognathi, Dipnoi, ...[thru] ...
Leptocardii, in the British Museum. Catalogue of Fishes i-xxv, 1-549.

Jiménez, S., Schönhuth, S., Lozano, I.J., González, J.A., Sevilla, R.G., Diez, A., Bautista,
J.M., 2007. Morphological, ecological, and molecular analyses separate *Muraena
augusti* from *Muraena helena* as a valid species. Copeia, 101-113.

Kass, R.E., Raftery, A.E., 1995. Bayes Factors. Journal of the American Statistical
Association 90, 773-795.

Lessios, H.A., 2008. The Great American Schism: divergence of marine organisms after the rise of the Central American Isthmus. Annual Review of Ecology, Evolution and Systematics 39, 63-91.

Loh, K.-H., Chen, I.-S., Randall, J.E., Chen, H.-M., 2008. A review and molecular phylogeny of the moray eel subfamily Uropterygiinae (Anguilliformes: Muraenidae) from Taiwan, with description of a new species. Raffles Bulletin of Zoology 19, 135-150.

Marko, P.B., 2002. Fossil calibration of molecular clocks and divergence times of geminate species pairs separated by the Isthmus of Panama. Molecular Biology and Evolution 19, 2005-2021.

Mehta, R.S., 2008. Ecomorphology of the moray bite: relationship between dietary extremes and morphological diversity. Physiological and Biochemical Zoology 82, 90-103.

Muss, A., Robertson, D.R., Stepien, C.A., Wirtz, P., Bowen, B.W., 2001. Phylogeography of *Ophioblennius*: The role of ocean currents and geography in reef fish evolution. Evolution 55, 561-572.

Myers, R.F., 1991. Micronesian reef fishes (2nd Edition). Coral Graphics, Guam.

Nelson, G.J., 1966. Gill arches of Teleostean fishes of the order Anguilliformes. Pacific Science 20, 391-408.

Peeters, F.J.C., Acheson, R., Brummer, G.J.A., De Ruijter, W.P.M., Schneider, R.R., Ganssen, G.M., Ufkes, E., Kroon, D., 2004. Vigorous exchange between the Indian and Atlantic oceans at the end of the past five glacial periods. Nature 430, 661-665.

Posada, D., 2008. jModelTest: Phylogenetic model averaging. Molecular Biology and Evolution 25, 1253-1256.

Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. Available from http://beast.bio.ed.ac.uk/Tracer.

Randall, J.E., 2007. Reef and shore fishes of the Hawaiian Islands. Sea Grant College Program, University of Hawaii, Honolulu.

Randall, J.E., Golani, D., 1995. Review of the moray eels (Anguilliformes: Muraenidae) of the Red Sea. Bulletin of Marine Science 56, 849-880.

Reece, J.S., Bowen, B.W., Joshi, K., Goz, V., Larson, A., 2010. Phylogeography of two moray eels indicates high dispersal throughout the Indo-Pacific. Journal of Heredity 102, 1-12.

Rocha, L.A., Bowen, B.W., 2008. Speciation in coral-reef fishes. Journal of Fish Biology 72, 1101-1121.

Rocha, L.A., Robertson, D.R., Rocha, C.R., Tassell, J.L., Craig, M.T., Bowen, B.W., 2005. Recent invasion of the tropical Atlantic by an Indo-Pacific coral reef fish. Molecular Ecology 14, 3921-3928.

Rocha, L.A., Rocha, C.R., Robertson, D.R., Bowen, B.W., 2008. Comparative phylogeography of Atlantic reef fishes indicates both origin and accumulation of diversity in the Caribbean. BMC Evolutionary Biology 8, 157.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Rozas, J.J., Sánchez-DelBarrio, C., Messeguer, X., Rozas, R., 2003. DNAsp, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496-2497.

Shannon, L.V., 1985. The Benguela ecosystem, I., evolution of the Benguela, physical features and processes. Oceanography and Marine Biology 23, 105-182.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16, 1114-1116.

Smith, A.G., Smith, D.G., Funnell, B.M., 2004. Atlas of Mesozoic and Cenozoic Coastlines. Cambridge University Press, New York.

Song, C.B., Near, T.J., Page, L.M., 1998. Phylogenetic relations among percid fishes as inferred from mitochondrial cytochrome b DNA sequence data. Molecular Phylogenetics and Evolution 10, 343-353.

Springer, V.G., 1982. Pacific Plate Biogeography, with special reference to shorefishes. Smithsonian Contributions to Zoology 367, 1-182.

Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. American Journal of Human Genetics 73, 1162-1169.

Stephens, M., Smith, N., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics 68, 978-989.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., Meyer, A., Bouvet, J., 1992. Unusually large mitochondrial variation in

populations of the blue tit, *Parus caeruleus*. Molecular Ecology 1, 27-36.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA

polymorphism. Genetics 123, 585-595.

Templeton, A.R., 1983. Phylogenetic inference from restriction endonuclease cleavage

site maps with particular reference to the evolution of humans and the apes.

Evolution 37, 221-244.

Veron, J.E.N., 1995. Corals in space and time. Cornell University Press, Ithaca.

Veron, J.E.N., Devantier, L.M., Turak, E., Green, A.L., Kininmonth, S., Stafford-Smith,

M., Peterson, N., 2009. Delineating the Coral Triangle. Galaxea, Journal of Coral

Reef Studies 11, 91-100.

Westneat, M.W., Alfaro, M.E., 2005. Phylogenetic relationships and evolutionary history

of the reef fish family Labridae. Molecular Phylogenetics and Evolution 36, 370-

390.

CHAPTER 3

PAGES 32-72

PHYLOGEOGRAPHY OF TWO MORAY EELS (*GYMNOTHORAX FLAVIMARGINATUS, G. UNDULATUS*) INDICATES HIGH DISPERSAL THROUGHOUT THE INDO-PACIFIC

Introduction

Coral reefs occur sporadically in a matrix of open ocean that spans 70% of the planet. Reef inhabitants that disperse only as adults can show population genetic structuring on scales of tens of kilometers (Bernardi, 2000). However, a pelagic (oceanic) larval stage permits most reef fishes to disperse across open ocean. Larvae either swim actively or drift passively in surface currents, sometimes modifying their buoyancy to use both surface and deep currents (Atema et al., 2002; Kingsford et al., 2002; Leis, 1991; Montgomery et al., 2001). This ability to disperse across unsuitable habitat prior to settling on a suitable reef enables dispersal among regions separated by thousands of kilometers.

The geographic distance over which larvae disperse varies across taxa and geographic regions, and is probably influenced by the amount of time spent in the pelagic stage (the pelagic larval duration, PLD) (Crisp, 1978; Ekman, 1953; Scheltema, 1968). The simplest expectation is that long PLDs yield proportionately greater gene flow among populations (Jones et al., 2005; Taylor and Hellberg, 2005). For many reef fishes, this prediction is upheld by population genetic patterns (Craig et al., 2007). However, PLD is not necessarily predictive (Weersing and Toonen 2009), and many reef species maintain greater population genetic structure than expected from their PLD (Bernardi et al., 2003; Planes and Fauvelot, 2002; Swearer et al., 2002; Taylor and Hellberg, 2003; Thacker et al., 2007). These findings have prompted research into the types of geographic, oceanographic, and biological factors that mediate larval dispersal, including variation in effective population size, habitat preference, and demographic changes in

response to major climatic and environmental shifts (Floeter et al., 2008; Hellberg, 2007; Leis, 2007; Leis et al., 2007; Rocha et al., 2007).

Genetic fragmentation among populations of reef fish species often coincides with major oceanic and geographic barriers (Briggs, 1961; Rocha et al., 2007; Springer, 1982). The two major phylogeographic breaks of the Indo-Pacific occur across the Eastern Pacific Barrier (EPB; 5,000 to 7,000 kilometers of open ocean that separate eastern and central Pacific reefs) and the Sunda Shelf (SS; an area of shallow reefs, located just west of Indonesia, that are exposed during low sea levels and separate the Indian and Pacific Oceans; Figure 3-1).

Some reef fishes show high genetic connectivity across thousands of kilometers (Muss et al., 2001; Rocha et al., 2002; Rocha et al., 2007), indicating gene flow across the Sunda Shelf (Horne et al., 2008) or the Eastern Pacific Barrier (Lessios and Robertson, 2006). However, no reef fish with a non-pelagic adult stage has been shown to maintain connectivity across both the Sunda Shelf and the Eastern Pacific Barrier. Fish species showing geographically widespread populations typically have PLDs exceeding 45 days, whereas most reef fishes have PLDs between 15 and 30 days (Lester and Ruttenberg, 2005). Surgeonfishes of the genus *Naso* are among the most extreme examples of reef fishes maintaining geographically widespread gene flow. *Naso vlamingii*, (Klanten et al., 2007) and two congeners (Horne et al., 2008) demonstrate similarly high levels of mitochondrial haplotype diversity and lack of phylogeographic structure across their Indo-Pacific ranges west of the Eastern Pacific Barrier. These three species share extended PLDs and similar life histories.

Moray eels (Muraenidae) are solitary predators in reef ecosystems. They comprise

about 200 species globally and approximately 150 species within the Indo-Pacific. Many

moray species span the entire Indo-Pacific, occupying coral reefs and rocky ledges from

the intertidal zone to 200+ meters deep. Moray eels are poor swimmers as juveniles and

adults and maintain high site fidelity to a few square meters of reef (Böhlke et al. 1989).

In a pattern typical of reef fishes, dispersal occurs by pelagic larvae. However, the

morays and elopomorph fishes (which includes all true eels, tarpons, tenpounders and

ladyfish) are unique in having a slender elongate larval form called a *leptocephalus*

(Figure 3-1). These larvae are among the simplest, long-lived, and self-sustaining

vertebrate forms. They are transparent except for eye pigmentation, and the body wall

may be only a few cells thick. They have no identifiable digestive activity or nutrient

stores, yet leptocephalus larvae can persist up to two years in pelagic environments

sustained by consuming dissolved organic carbon and the fecal pellets and waste products

of zooplankton and other larvae (Böhlke et al., 1989; Mochioka and Iwamizu, 1996;

Ishikawa et al., 2001).

Moray eels are difficult to sample because of their solitary, aggressive, reclusive

habits, and this study represents the first phylogeographic survey of the group. Previous

authors have suggested that they have extensive oceanic dispersal, based on the extended

PLD of the leptocephalus larvae, and "the broad distributions of many species within the

vast Indo-Pacific region" (Randall 2007). Here I test this hypothesis for the first time

with phylogeographic surveys and inferred coalescence times for haplotype variation

within *Gymnothorax flavimarginatus* and *G. undulatus*. Both species are ecological

generalists on reefs from 0 to >150 meters deep, with distributions that traverse the Sunda

Shelf and the Eastern Pacific Barrier. I sample *G. undulatus* and *G. flavimarginatus* from reefs across the Indo-Pacific, from South Africa to Panama, using portions of the mitochondrial genes encoding cytochrome *b* (*CYB*) and cytochrome oxidase subunit 1 (*CO1*), plus portions of two nuclear recombination activation loci, *RAG-1* and *RAG-2*. My results reveal the first cases of high gene flow through larval dispersal across both the Sunda Shelf and the Eastern Pacific Barrier.

Methods

Specimens were captured in fish traps, lobster traps, and by pole spears while snorkeling or SCUBA diving. Here I report the collection of 289 specimens with only two injuries, a minor bite wound to the hand of collaborator Brian Bowen from *G. undulatus*, and a bite to the foot of documentary filmmaker Ziggy Livnat. This is a notable accomplishment, as moray eels are widely (and rightfully) feared for sudden and egregious acts of aggression (Riordan et al. 2004). In particular, *G. undulatus* is "especially vicious" (Hiatt and Strasburg 1960), and several bite victims are known to the author (see also Randall 2007).

Fin clips or muscle samples were taken from live specimens and other tissue samples from tissue-storage banks. Sampling includes 124 *G. flavimarginatus* specimens from 15 sites and 165 specimens of *G. undulatus* from 14 sites spanning the Indo-Pacific (Table 3-1, Figure 3-1). All samples were collected between 2001 and 2009. DNA was isolated using Viogene Genomic DNA Extraction column kits. Polymerase chain reactions were performed in 25μL reactions of 5μL of Promega (www.promega.com) 5x buffer, 2.5μL of 25mM $MgCl_2$, 2.5μL of 0.2μM dNTPs, 2.5μL of 0.2μM of each primer,

0.125µL (1 unit) of Promega GoTaq DNA polymerase, and 2µL of template DNA at approximately 5ng/µL. A 632-bp fragment of the gene encoding *cytochrome b* (*CYB*) was amplified using the primers L14725 (5'-GTG ACT TGA AAA ACC ACC GTT G-3') (Song et al., 1998) and H15573 (5'-AAT AGG AAG TAT CAT TCG GGT TTG ATG-3') (Taberlet et al., 1992) and an annealing temperature of 50°C. A 596-bp fragment of the gene encoding *cytochrome oxidase subunit 1* (*COI*) was amplified using primers FishF2 (5'-TCG ACT AAT CAT AAA GAT ATC GGC AC-3') and FishR2 (5'-ACT TCA GGG TGA CCG AAG AAT CAG AA-3') (Ward et al., 2005) and annealing temperature of 50°C. A 421-bp fragment of the nuclear recombination activation gene *RAG-1* was amplified using primers designed based on published marine fish *RAG-1* sequences: RAG1-F3 (5'-GCC TCA GAA AAC ATG GTG CT-3') and RAG1-R3 (5'-CCA CAC AGG TTT CAT CTG GA-3') with an annealing temperature of 50°C. A 754-bp fragment of the nuclear recombination activation gene *RAG-2* was sequenced using primers RAG2-F3 (5'-AGG TGA CCC TTC GTT GTC AG-3') and RAG2-R3 (5'-ATG AGG CTC CCT TCC AAA GT-3'), also designed, at an annealing temperature of 52°C. The thermal profiles for PCR were 95°C for 3 min, followed by 35 cycles of 95°C for 1 min, annealing temperature for 40 sec, and 72°C for 45 sec, with a final elongation at 72°C for 7 min.

PCR products were visualized through agarose-gel electrophoresis and purified using Exo-Sap or Viogene Gel Purification Kits (www.viogene.com). Sequences were generated on ABI 3130 and ABI 3330 automated DNA sequencers at the Washington University Genome Sequencing Center and the Smithsonian Museum Support Center. I sequenced all loci in both the forward and reverse directions, and I generated and edited

all of the sequence data. DNA sequences were manually edited using SEQUENCER 4.8 (Ann Arbor, MI) and aligned by hand. For nuclear markers, heterozygous positions were identified as cases wherein a secondary peak in the electropherograms reached at least 25% of the intensity of the primary peak. Gametic phases of nuclear sequences with more than a single heterozygous site were estimated using a Bayesian approach implemented in the software program PHASE 2.1 (Stephens and Donnelly, 2003; Stephens et al., 2001). All PHASE analyses were run through five iterations, each initiated using a different random-number seed and run for 1000 iterations with a single thinning interval and 100 burn-in iterations. Consistency of results was determined by examining allele frequencies and coalescent goodness-of-fit measures estimated for each of the four runs. In the case that the phase of a polymorphism could not be confidently estimated with 90% posterior probability, that site was coded as missing data. Less than 3% of all nucleotide characters were coded as missing data under this criterion. None of the markers for either species appear to be under selection according to Tajima's D (Tajima 1989) tests for selection (all $P > 0.1$) in ARLEQUIN 3.1 (Excoffier and Schneider, 2005). Lastly, tests of linkage disequilibrium in ARLEQUIN between the nuclear loci *RAG-1* and *RAG-2* were non-significant ($P > 0.1$) for both species.

I constructed haplotype networks for the mitochondrial fragments from *CYB* and *COI* concatenated as a single genetic marker, and separately for each of the nuclear markers *RAG-1* and *RAG-2*. In all cases, haplotype networks were constructed with TCS 2.1 (Clement et al. 2000) under 95% statistical parsimony criteria. This approach permits a qualitative assessment of the geographic distributions of alleles. The standard molecular

diversity measures of haplotype diversity (h) and nucleotide diversity (π) were calculated in ARLEQUIN according to Nei (1987).

I tested for population structure and phylogeographic patterns with three independent analyses: 1) an Analysis of Molecular Variance (AMOVA; Excoffier et al. 1992), 2) pairwise Φ_{st} and a Mantel test of geographic vs. genetic distances (Φ_{st}), and 3) permutation tests of geographic distances on haplotype networks. Sample sizes less than N=5 were excluded from any pairwise comparisons. After preliminary trials, the two mtDNA fragments were analyzed as a single composite sequence. I subjected each marker (mtDNA, *RAG-1, RAG-2*) to AMOVA using ARLEQUIN based on haplotype frequencies and molecular distances among haplotypes. Individuals were grouped by localities and then by three larger subgroups comprising localities from the Indian Ocean, the western and central Pacific Ocean, and the eastern Pacific Ocean. These partitions correspond to the three oceanic regions demarcated by the Sunda Shelf and the Eastern Pacific Barrier (Figure 3-1). Statistical significance of the corresponding ARLEQUIN statistic, Φ_{sc}, was determined by 10,000 permutations.

I conducted Mantel tests in the program PASSAGE 2.0 (Rosenberg, 2001) to measure correlations between geographic and genetic distances among populations. Genetic distances for all loci were computed in ARLEQUIN as pairwise Φ_{st} values based on the number of nucleotide differences between alleles and allele frequency differences between populations. I report only uncorrected differences because corrected pairwise differences under a variety of mutation models yielded congruent results. Significance of pairwise Φ_{st} values was computed using the B-H (Benjamini and Hochberg 1995; Misawa et al. 2008) and the more conservative Bonferroni (Bonferroni 1936) corrections

for multiple comparisons. The pairwise matrix of geographic distances between populations was obtained with ARCGIS 9.3 as the shortest straight-line distance that did not cross dry land. I could not incorporate oceanic currents in this analysis because most available data are for surface currents, whereas moray eel larvae utilize both surface currents and counter-flowing deeper currents (Böhlke et al., 1989). Geographic and genetic distance (Φ_{st}) matrices were tested for correlation separately for each molecular maker (mtDNA, *RAG-1, RAG-2*).

Associations of geographic and genetic distances were resolved using a permuted chi-squared test of genetic distances among hierarchically nested groups of haplotypes and geographic areas covered by distributions of grouped haplotypes (Templeton, 1998). This approach is the initial step in nested clade phylogeographic analysis (NCPA; Templeton 1998) and mirrors the Mantel test, but instead of using single populations as the sole level of sampling, it tests for associations between geographic distance and hierarchically nested groups from the haplotype network. Geographic distances were identical to those of the Mantel tests. Tests were performed separately on each locus (mtDNA, *RAG-1, RAG-2*) in the program GEODIS 2.6 (which includes the Dunn-Sidak correction for multiple tests; (Posada et al., 2000) using 10,000 permutations. Inferences were made using the 2008 inference key available at http://darwin.uvigo.es/software/geodis.html. Loops denoting ambiguous groupings in haplotype networks were treated as follows: all possible nested combinations of resolutions to the loops were computed, and analyses were repeated for each possible network configuration. Alternative resolutions of loops did not alter inferences in this study.

There is little information available on the time since divergence from a most recent common ancestor of moray eels and their closest relatives, the genera *Conger* and *Anguilla* (Inoue et al., 2003). Without a strong fossil record or obvious cases of vicariance, one cannot derive a species-specific mutation rate for estimating coalescence times. For the mitochondrial gene *CYB,* I used a mutation rate of 2% per MY, which has been widely applied to phylogeographic studies of reef fishes (Bermingham et al., 1997; Bowen et al., 2001; Bowen et al., 2006b) and is supported by measurements for goby (*Gnatholepis*) species separated by the Isthmus of Panama (1.95-2.17% divergence per MY; Rocha et al. 2005). I used a mutation rate for the *COI* data of 1.2% sequence divergence per MY as calibrated by Bermingham et al. (1997) for marine fishes. I estimated the time to coalescence for standing genetic variation in each of the two study species, and for haplotype groups with significant inferences in NCPA, using the Bayesian Markov chain Monte Carlo (MCMC) approach as implemented in BEAST 1.5.2 (Drummond and Rambaut 2007). I used a congeneric outgroup and determined the target species (*G. undulatus* or *G. flavimarginatus*) to be monophyletic for haplotypic variation. Models of evolution were computed in MrModelTest (Posada and Buckley 2004) and the combined mtDNA data were partitioned by gene region by codon position, for a total of six partitions. A fixed mutation rate was estimated at 1.6% divergence per MY (the weighted average of 2% and 1.2% for *CYB* and *COI*, respectively), and input as 0.008 substitutions per site per lineage per MY. A coalescent tree prior of expansion growth was used, as recommended for estimating intraspecific coalescence times when population growth is inferred (see results of mismatched distributions) (Drummond and Rambaut 2007). Simulations ran for 30 million generations, with sampling every 3000

generations. Two independent runs were computed for each species, and runs were then

combined with a 10% burnin using LogCombiner 1.5.2 (Drummond and Rambaut 2007).

Convergence of chains and parameter estimates in the MCMC run were computed in the

program Tracer 1.4 (Rambaut and Drummond 2007). The Effective Sample Size (ESS)

values, a measure of how well parameter space is evaluated, for each parameter were

over the recommended value of 200 (Drummond and Rambaut 2007), and 95% highest

posterior density intervals (HPD) are reported as encompassing the uncertainty around

coalescence time estimates within each species. This estimate of confidence intervals is

directly proportional to and does not account for variation in the mutation rates used.

Frequency distributions of numbers of mutational differences between haplotypes

sampled within a species (= mismatch distribution) were computed in ARLEQUIN using

the least-square method (Schneider et al., 2000). To compute the confidence intervals

around these values and the ability to reject a model of demographic expansion through

time, I performed 100,000 parametric bootstraps and calculated Harpending's

Raggedness Index (HRI) (Harpending 1994). I also computed Fu's F_S statistics (Fu 1997)

to identify cases of demographic expansion, which are indicated by significant negative

values.

Results

Nuclear and mitochondrial haplotype networks for *Gymnothorax flavimarginatus*

and *G. undulatus* are presented in Figure 3-2. Notably, haplotypes from the major

geographic subdivisions of the Indo-Pacific are dispersed throughout the networks.

Combining *CYB* and *COI*, I identify 104 mtDNA haplotypes for *G. flavimarginatus*

(N=124) (GenBank Accession GU175445-GU175541) and 112 haplotypes for *G. undulatus* (N = 165) (GU175576-GU175657). No mitochondrial haplotype occurs in samples of all three major oceanic regions (Indian Ocean, western/central Pacific Ocean, eastern Pacific Ocean), but four haplotypes in *G. flavimarginatus* and six in *G. undulatus* occur in adjacent regions; all other haplotypes are observed in a single region. The nuclear markers *RAG-1* and *RAG-2* reveal less variation than the combined mitochondrial markers. *RAG-1* sequences resolved 16 haplotypes for *G. flavimarginatus* (GU175542-GU175557) and 23 haplotypes for *G. undulatus* (GU175658-GU175680). *RAG-2* sequences included 18 haplotypes for *G. flavimarginatus* (GU175558-GU175575) and 22 haplotypes for *G. undulatus* (GU175681-GU175702). The most common nuclear haplotypes occur in most localities, with haplotype sharing among the three major oceanic regions (Figure 3-2). Groupings of haplotypes endemic to one of the three major oceanic regions occur only in the mitochondrial haplotype networks. Molecular diversity indices for *G. flavimarginatus* at mtDNA were $h = 0.997$ and $\pi = 0.012$; at *RAG-1* $h = 0.706$ and $\pi = 0.0050$, and at *RAG-2* $h = 0.912$ and $\pi = 0.0043$. Molecular diversity indices for *G. undulatus* at mtDNA were $h = 0.972$ and $\pi = 0.012$; at *RAG-1* $h = 0.808$ and $\pi = 0.0053$, and at *RAG-2* $h = 0.907$ and $\pi = 0.0067$.

Results of a locus-by-locus AMOVA revealed no population structure among the three geographic groupings demarcated by the Eastern Pacific Barrier and the Sunda Shelf. For *G. flavimarginatus*, the mtDNA yields a non-significant $\Phi_{sc} = 0.0088$ ($P = 0.29$) and the nuclear markers a non-significant $\Phi_{sc} < 0.00001$ ($P = 0.89$) for *RAG-1* and $\Phi_{sc} < 0.00001$ ($P = 0.73$) for *RAG-2*. The mtDNA results for *G. undulatus* reveal a low but significant $\Phi_{sc} = 0.039$ ($P = 0.007$). The nuclear markers yield non-significant $\Phi_{sc} <$

0.00001 ($P = 0.30$) for *RAG-1* and $\Phi_{sc} < 0.00001$ ($P = 0.99$) for *RAG-2*. Results for both

species identify the vast majority of the variation (95-100%) occurring within

populations.

The Mantel tests for correlations between geographic distances and Φ_{st} values are

not significant for mtDNA, *RAG-1*, and *RAG-2* (all $R < 0.0001$, *P*-values greater than

0.2). Tables 3-2 and 3-3 present mtDNA pairwise Φ_{st} values for *G. flavimarginatus* and

G. undulatus respectively. Structure in *Gymnothorax flavimarginatus* was especially low

with zero out of 55 significant comparisons for mtDNA (Table 3-2), and no significant

comparisons for *RAG-1* or *RAG-2*. Similarly, *Gymnothorax undulatus* showed low levels

of structure with 4 out of 36 significant comparisons for mtDNA (Table 3-3), and none

for *RAG-1* or *RAG-2*.

The NCPA of the combined mtDNA sequences in *G. flavimarginatus* revealed

only three nested groups that refute the hypothesis of panmixia. None of the resolutions

of loops altered the chain of inference or the significance of pairwise comparisons; Figure

3-2 shows resolved loops only for clades with significant inferences. Group A (Figure 3-

2; Table 3-4) supports an inference of contiguous range expansion, estimated to have

occurred within the last 0.6 MY (95% HPD 0.3 to 0.9 MY). Tip haplotype groups in

Group A show expansion from the internal and mostly Hawaiian group to tip groups with

haplotypes in the Seychelles, South Africa, and the eastern Pacific. Groups B and C yield

inferences of restricted gene flow with isolation by distance coalescing within the last 0.5

MY (95% HPD 0.2 to 0.8 MY) and 0.4 MY (95% HPD 0.3 to 0.7 MY), respectively.

Within group B, tip subgroup B-2 is observed only in Hawaii, whereas the more interior

subgroup B-1 from which it derives includes haplotypes distributed across the Indo-

Pacific. Both subgroups of group C include haplotypes from Hawaii, but they differ in having haplotypes from the Indian Ocean (subgroup C-2) versus the eastern Pacific Ocean (subgroup C-1).

The NCPA for *G. undulatus* includes two nesting levels (groups D and E) that refute the null model of panmixia, both involving range expansions of haplotypes (Figure 3-2; Table 3-4) that coalesce within 0.6 MY (95% HPD 0.4 to 0.8 MY) and 0.3 MY (95% HPD 0.2 to 0.5 MY), respectively. Range expansion in group D includes tip haplotypes sampled from Australia (subgroup D-1) or Fiji (subgroup D-2) from an inferred ancestral range in Hawaii, indicating some historical fragmentation followed by range expansion within the western Pacific. Group E shows a tip haplotype observed in Hawaii and the Indian Ocean (subgroup E-1) derived from an ancestral distribution in the Pacific Ocean (subgroup E-3).

The coalescence times for standing genetic variation based on mtDNA were 2.3 MY (95% HPD 1.7 to 3.0 MY) for *G. flavimarginatus* and 5.9 MY (95% HPD 3.5 to 8.6 MY) for *G. undulatus.* Estimates of coalescence times for *G. flavimarginatus* and *G. undulatus* are younger than those for other reef fishes with similar population structure (i.e., genus *Naso*; Klanten et al. 2007; Horne et al. 2008), but the confidence intervals surrounding those estimates broadly overlap.

I evaluated evidence for population expansion using mismatch distributions and Fu's F_S. Mismatch distributions test the null hypothesis of demographic expansion, while Fu's F_S tests the null hypothesis of constant population size, with significantly negative values indicating demographic expansion. Therefore, a non-significant result for the

mismatch distribution test and a significantly negative Fu's F_S both indicate population expansion. The mismatch distributions for mtDNA are shown in Figure 3-3. Overall, mismatched analyses indicate a pattern of population growth in both species for all markers. In *G. flavimarginatus*, mismatch analyses fail to reject the null hypothesis of demographic expansion for mtDNA ($P = 0.78$, HRI = 0.10; $P = 0.72$), *RAG-1* ($P = 0.10$, HRI = 0.04; $P = 0.55$), and *RAG-2* ($P = 0.22$, HRI = 0.28; $P = 0.52$). Similar results were obtained in mismatched analyses of *G. undulatus* for mtDNA ($P = 0.42$, HRI = 0.017; $P = 0.51$), *RAG-1* ($P = 0.5$, HRI = 0.03; $P = 0.68$), and *RAG-2* ($P = 0.99$, HRI = 0.004; $P = 0.99$). The results of the Fu's F_S analyses support these conclusions, with significant negative values in *G. flavimarginatus* for mtDNA (-24.0; $P < 0.001$) and *RAG-1* (-14.8; $P < 0.001$), but a non-significantly negative value for *RAG-2* (-3.2; $P = 0.15$). Fu's F_S analyses in *G. undulatus* indicated demographic expansion with significant negative values for mtDNA (-24.0; $P < 0.001$) and *RAG-1* (-7.2; $P = 0.019$), but with a non-significant value for *RAG-2* (0.57; $P = 0.65$). In summary, all analyses support conclusions of demographic expansion except Fu's F_S for *RAG-2* in both species. Peak pairwise differences of approximately 10 substitutions (*G. undulatus*) versus 17 substitutions (*G. flavimarginatus*) between paired haplotypes for the 1228 bp of mtDNA are consistent with transitory fragmentation of populations between 300,000-600,000 years BP (the estimated coalescent times for clades with significant inferences in NCPA). These events yielded divergent haplotype lineages that have persisted within each species through subsequent climatic cycles of approximately 100,000 years (Paillard 2006).

Discussion

Several reef fishes are known to have highly dispersive pelagic stages, including soldierfishes (genus *Myripristis*; Bowen et al. 2006b; Craig et al. 2007), pygmy angelfishes (genus *Centropyge*; Bowen et al. 2006a; Schultz et al. 2007), and unicornfishes (genus *Naso*; Klanten et al. 2007; Horne et al. 2008); however, even these species show genetic partitions (or distribution limits) at major biogeographic barriers. In the first moray species subject to phylogeographic analyses, I reveal through three independent tests a prevailing pattern of very little to no genetic structure among populations separated by the major biogeographic barriers of the eastern Pacific and Sunda Shelf. Leptocephalus larvae in general are extremely long-lived in the pelagic environment and can delay metamorphosis until appropriate conditions are available (Castonguay, 1987; Crabtree et al., 1992; Schmidt, 1923; Tseng, 1990), and it is likely that these capabilities have enabled the high levels of connectivity that I observe throughout the Indo-Pacific.

Haplotypes from the Indian, western/central Pacific, and eastern Pacific Oceans are interspersed on the networks for all three loci (mtDNA, *RAG-1, RAG-2*) with geographic clustering of haplotypes occurring only near some tips of the mtDNA network (Figure 3-2). Nuclear haplotypes are shared among localities separated by more than 22,000 kilometers and across two major biogeographic barriers. This high genetic continuity among populations with low adult vagility strongly implicates the leptocephalus larva (especially the long pelagic stage) as the dispersal engine.

The AMOVA results indicate a lack of persistent biogeographic barriers to gene flow. Only the mitochondrial locus of *G. undulatus* reveals a significant Φ_{sc} among

regions delineated by the SS and EPB, with biogeographic barriers explaining only 3.9%

of the overall variation. Mantel tests of genetic and geographic distances likewise fail to

reject panmixia among geographic populations. Although the SS and EPB represent

species boundaries in many taxa (Briggs, 1961), and substantial obstacles to gene flow

within species that cross them (Randall, 1998), they appear porous to gene flow among

populations of moray eels. Notably, the pairwise Φ_{st} comparisons show no consistent

breaks across the SS, but *G. undulatus* has low/significant structure across the EPB in

two comparisons with Hawaiian islands. I conclude that the oceanic EPB can be a

substantial barrier under contemporary conditions, whereas the SS seems to be important

in a historical context but not a barrier under contemporary conditions. Leptocephalus

dispersal capabilities quickly erase genetic separations developed across the SS during

glacial maxima, while the signature patterns of partial isolation across the EPB are

ongoing and detectable at low levels. While sampling error should increase the variance

on estimates of genetic divergence relative to actual divergence, my finding of

widespread genetic homogeneity among populations is robust to the relatively small

sample sizes within populations of the Indian and eastern Pacific Oceans (Table 3-1).

My NCPA results for the mitochondrial DNA indicate a pattern of transitory

fragmentation among populations followed by genetic merging within 0.3 to 0.6 MY.

The NCPA analysis is sensitive to historical events such as range expansion and previous

fragmentation, and the mitochondrial locus is the only one variable enough to reveal such

patterns. Three groupings of mitochondrial haplotypes from *G. flavimarginatus,* each

group estimated to be less than 0.6 MY old, show evidence of geographic fragmentation.

Within group A (Figure 3-2, Table 3-4), mitochondrial haplotypes of *G. flavimarginatus*

show evidence of a range expansion across the Sunda Shelf within the past approximately

0.6 MY. Haplotypes in groups B and C also show some evidence of fragmentation on the

timescale of 0.5 to 0.3 MY, respectively. Within *G. undulatus*, mtDNA haplotype groups

D and E show evidence of range expansion on this same timescale across the Sunda Shelf

(group E) and between Australian and Hawaiian regions of the western/central Pacific

Ocean. Results of the mismatch distributions and occurrence of four low but statistically

nonzero pairwise Φ_{st} values for the mitochondrial markers of *G. undulatus* (Table 3-3)

are consistent with this hypothesis of transitory geographic genetic structure within the

past 0.6 MY. Based on the estimated coalescence times within the past 2.3 MY in *G.

flavimarginatus* and 5.9 MY in *G. undulatus*, the prevailing pattern seems to be genetic

homogeneity of populations over the longer (million year) timeframe, with transitory

population fragmentation on the shorter timescale of 0.3 to 0.6 MY.

Recent controversy regarding NCPA (Garrick et al., 2008; Knowles, 2008; Petit,

2008) concerns whether significant genetic associations of haplotype groupings with

geographic distances justify inferences of isolation by distance or contiguous range

expansion, as these are the two most common inferences from datasets generated under

panmictic conditions by Panchal and Beaumont (2007). It should be noted that the

limitations described by Petit (2008) and Panchal and Beaumont (2007) have been

corrected (Templeton 2008, 2009). My results from NCPA involve three and four-step

clades with significant non-random associations of genetic and geographic distances. My

interpretations from NCPA are consistent with the development of transient,

geographically localized population genetic structure on the order of a few hundred

thousand years, and these are supported in part by significant pairwise Φ_{ST} comparisons

and mismatch distributions. At broader scales and the entire cladogram level, NCPA, AMOVA, Mantel Tests, and pairwise Φ_{ST} all yield the same result: inability to reject panmixia. At best, NCPA has identified evidence for population genetic structure at a level that is lost to more reductionist approaches such as AMOVA. At worst, it has yielded a false inference, and all populations of *G. undulatus* and *G. flavimarginatus* surveyed here are panmictic. In either case, haplotypes likely to have arisen within the past few hundred thousands of years have not all acquired ocean-wide distributions, but genetic continuity of populations is nonetheless the prevailing pattern when results are averaged over the multi-million year history covered by mtDNA markers.

Horne et al. (2008) and Klanten et al. (2007) apply the term "temporal genetic partitioning" to Indo-Pacific reef fishes to describe a pattern of population genetic homogeneity on a million-year time scale with transitory fragmentation and demographic expansion on shorter time scales. "Each expansion episode is likely to have been accompanied by re-colonization of distant reef habitats resulting in repeated and widespread secondary contacts amongst previously isolated, genetically divergent demes" (Klanten et al. 2007). Global climatic fluctuations provide the most obvious source of recurring cycles of fragmentation followed by expansion of populations for reef fishes having widespread larval dispersal. Randall (1998) hypothesizes three to six episodes within the past 700,000 years during which sea level would have been low enough to preclude dispersal of marine species between the Indian and Pacific Oceans. A review of geological climatic data by Paillard (2006) indicates sea-level fluctuation occurring on a cycle of ~100,000 years throughout the past million years. These results are highly congruent with my conclusion that the two *Gymnothorax* species are emerging

from a transitory fragmentation event with some remnants of fragmentation and recent range expansions over the past 0.3 to 0.6 MY shown by some of the younger mitochondrial haplotypes.

This interpretation is consistent with evidence for population growth in the form of mismatched distributions and significant negative Fu's F_S values for the mitochondrial loci of each species, and for *RAG-1* of *G. undulatus*. As proposed for other reef fishes and invertebrates, this population expansion likely followed the last period of low sea levels that exposed portions of the Sunda Shelf (Barber et al. 2006; Klanten et al. 2007; Gaither et al. 2010).

The Hawaiian Archipelago has a special biogeographic significance for moray eels, and this is apparent in several facets of my data. First, morays are common and abundant on these islands, one of the most isolated tropical archipelagos in the world, while common reef predators such as groupers (Serranidae) and snappers (Lutjanidae) are nearly absent. Thus the high dispersal evident in this survey has allowed the morays to colonize and to proliferate in Hawaii, to the exclusion of many west/south Pacific competitors (Randall 2007). Second, Hawaii may be the biogeographic crossroads between the tropical eastern Pacific and western Pacific. This is indicated by the distribution of shared haplotypes from Hawaii to the western Pacific and Indian Ocean (subgroup C-2 and group B in *G. flavimarginatus*, subgroup E-1 in *G. undulatus*), and Hawaii to the eastern Pacific (subgroups A-1 and C-1 in *G. flavimarginatus*; Figure 3-2; Table 3-4). Based on the latter network configuration, NCPA reveals dispersal from Hawaii to the eastern Pacific in the last 0.3 MY (directionality inferred from haplotype network rooting and historical biogeography; see Briggs 1999). Many tropical species

show very shallow phylogeographic histories in the eastern Pacific, indicating that much of this fauna is extirpated during glacial conditions (Bowen and Karl 2007). Hence, Hawaii may be the essential source of moray diversity in the East Pacific. Third, the only significant mtDNA partitions on a local scale for either species are within the Hawaiian Islands, and one of them is between the high islands of the southeastern Hawaiian Archipelago and the low islands to the northwest. The latter region is now protected as the Papahānaumokuākea Marine National Monument. Emerging data indicate that the jurisdictional and ecological differences between these regions are matched by population genetic differences in some other reef species (Rivera et al. 2004; Ramon et al. 2008; Eble et al. 2009).

Gymnothorax flavimarginatus and *G. undulatus* represent extreme cases among reef fishes in maintaining gene flow across two thirds of the planet. Several other fishes found on or near coral reefs maintain high levels of connectivity across similar spatial scales (e.g., whale sharks, Castro et al. 2007; tunas, Theisen et al. 2008), but these all have highly vagile adult forms that are not restricted to coral-reef habitat. My favored hypothesis for the geographic-genetic homogeneity of these *Gymnothorax* species is that their extended pelagic larval phase permits widespread dispersal of individuals following the removal of transitory oceanic barriers to pelagic dispersal. The genetic connectivity of both *Gymnothorax* species across the Eastern Pacific Barrier and Sunda Shelf indicates an ability to disperse across even the strongest obstacles to other reef fishes.

Table 3-1. Sample sizes at each location (with zeros indicated by a "-"), subdivided into three regions corresponding to the biogeographic barriers of the Sunda Shelf and the Eastern Pacific Barrier. NWHI= Northwestern Hawaiian Islands, MHI= Main Hawaiian Islands

		G. flavimarginatus	*G. undulatus*	Latitude	Longitude
Indian Ocean					
1	South Africa	7	6	33°44'S	26°37'E
2	Seychelles	5	7	8°21'S	50°31'E
Western/Central Pacific					
3	South China Sea	1	3	18°40'N	118°45'W
4	Philippines	6	-	11°35'N	123°49'W
5	Saipan	5	-	15°7'N	145°43'W
6	Australia	7	7	24°41'S	152°42'E
7	Midway (NWHI)	-	3	28°12'N	177°19'W
8	Pearl & Hermes (NWHI)	3	3	27°50'N	175°50'W
9	Laysan (NWHI)	5	-	25°47'N	171°45'W
10	Maro (NWHI)	10	29	25°37'N	170°33'W
11	Gardner Pinnacles (NWHI)	-	1	25°0'N	167°55'W
12	Necker (NWHI)	3	34	23°36'N	164°39'W
13	Oahu (MHI)	7	19	21°30'N	157°45'W
14	Hawaii (MHI)	53	27	19°5'N	155°29'W
15	Fiji	1	5	17°39'S	177°7'E
16	American Samoa	5	-	14°12'S	170°55'W
17	Moorea	-	2	17°36'S	149°38'W
Eastern Pacific					
18	Clipperton	6	-	10°22'N	109°31'W
19	Panama	-	19	8°47'N	79°0'W
	Total	124	165		

Table 3-2. Pairwise Φ_{st} values calculated from mtDNA haplotypes (below diagonal) and the mean value for nuclear loci *RAG-1* and *RAG-2* (above diagonal) for *Gymnothorax flavimarginatus*. None of the pairwise comparisons were significantly greater than zero for mtDNA, *RAG-1*, or *RAG-2* after B-H or Bonferroni corrections for multiple comparisons. Populations are numbered as in Figure 3-1 and Table 3-1. Values less than zero have no biological meaning and should be considered indistinguishable from zero.

	1	2	4	5	6	9	10	13	14	16	18
1	------	-0.05	-0.07	-0.09	-0.01	-0.02	-0.06	-0.01	-0.03	-0.07	-0.04
2	-0.07	------	-0.06	-0.08	-0.01	-0.02	-0.03	-0.02	0.01	-0.06	0.01
4	-0.05	-0.12	------	-0.09	0.01	-0.02	-0.04	-0.03	-0.03	-0.07	-0.03
5	-0.02	-0.03	-0.02	------	0.01	-0.01	-0.05	-0.04	-0.02	-0.07	-0.04
6	-0.08	-0.10	-0.07	-0.04	------	-0.06	-0.03	0.09	0.04	0.01	0.04
9	-0.05	-0.10	-0.09	-0.03	-0.03	------	-0.05	0.06	0.02	-0.09	0.01
10	-0.01	-0.06	0.00	0.01	0.02	-0.07	------	0.0	-0.03	-0.06	0.02
13	-0.06	-0.05	-0.05	-0.07	-0.06	-0.04	0.03	------	-0.01	0.0	0.04
14	-0.03	-0.05	0.00	0.01	-0.01	-0.02	0.02	-0.01	------	-0.03	-0.04
16	-0.04	-0.07	-0.07	-0.01	-0.08	-0.01	0.09	-0.02	0.02	------	-0.07
18	-0.04	-0.11	-0.08	-0.04	-0.06	-0.08	-0.05	-0.01	0.00	-0.05	------

Table 3-3. Pairwise Φ_{st} values calculated from mtDNA haplotypes (below diagonal) and the mean value for nuclear loci *RAG-1* and *RAG-2* (above diagonal) for *Gymnothorax undulatus*, with values significantly greater than zero in bold type. Populations are numbered as in Figure 3-1 and Table 3-1. The mtDNA comparison between Maro Island (#10) and Hawaii (#14) was significant after the B-H correction, but not after the more conservative Bonferroni correction. None of the pairwise comparisons for *RAG-1* or *RAG-2* were significantly different from zero after B-H or Bonferroni corrections for multiple comparisons. Values less than zero have no biological meaning and should be considered indistinguishable from zero.

	1	2	6	10	12	13	14	15	19
1	------	-0.06	-0.05	-0.04	-0.03	0.04	0.02	-0.04	-0.04
2	0.01	------	-0.05	-0.02	-0.03	0.05	0.02	-0.06	-0.03
6	0.08	-0.05	------	-0.05	-0.05	-0.01	0.0	-0.05	-0.04
10	0.06	0.00	0.00	------	-0.03	-0.01	-0.01	-0.02	-0.03
12	-0.03	0.07	0.08	0.04	------	0.01	-0.01	-0.03	-0.01
13	0.04	-0.02	-0.03	0.00	0.05	------	-0.03	0.07	0.02
14	0.03	0.11	0.14	**0.06**	-0.01	**0.10**	------	0.04	0.01
15	-0.05	-0.03	-0.02	-0.02	-0.02	-0.02	-0.01	------	0.03
19	0.03	0.00	-0.05	0.06	**0.09**	0.05	**0.13**	0.01	------

Table 3-4. Tests for geographically non-random distributions of nested groupings of

mitochondrial haplotypes (NCPA; Templeton, 1998) using the December 2008 inference

key. Groups are listed as marked on Figure 3-2. All statistically significant departures

from random distributions indicate range expansions or geographic fragmentation among

relatively young haplotype groups (within the past 0.3 to 0.6 MY). No significant

departures from randomness were detected for nuclear markers in either species or for

older groupings of mitochondrial haplotypes. Also reported are coalescence times (and

95% HPD intervals) in millions of years (MY) as estimated under a coalescent model of

demographic expansion in the program BEAST.

Group	Chain of inference	Inference	Coalescence Time
		G. flavimarginatus	
Group A	1-2-11-12-No	Contiguous range expansion	0.6 (0.3-0.9) MY
Group B	1-2-3-4-No	Restricted gene flow with IBD	0.5 (0.2-0.8) MY
Group C	1-2-3-4-No	Restricted gene flow with IBD	0.4 (0.3-0.7) MY
		G. undulatus	
Group D	1-2-11-12-No	Contiguous range expansion	0.6 (0.4-0.8) MY
Group E	1-2-11-12-No	Contiguous range expansion	0.3 (0.2-0.5) MY

Figure Legends

Figure 3-1. Map of sampling localities and sketch of generalized moray leptocephalus larvae (bottom right). Each sampling locality is indicated by a number in parentheses that corresponds to localities in Figure 3-2 and Tables 3-1 through 3-3. Localities are categorized as Indian Ocean (black shading), western/central Pacific Ocean (open circles) or eastern Pacific Ocean (gray shading) to denote the three major subdivisions of the Indo-Pacific. See Table 3-1 for sample size at each locality. Dotted lines represent the biogeographic barriers of the Sunda Shelf (SS) and the Eastern Pacific Barrier (EPB).

Figure 3-2. Haplotype networks for mitochondrial DNA (left) and for *RAG-1* and *RAG-2* for both *Gymnothorax flavimarginatus* and *Gymnothorax undulatus*. Rectangles represent inferred ancestral haplotypes, circles represent additional haplotypes, and small black-filled circles represent inferred missing haplotypes needed to connect observed haplotypes. The size of the circle or rectangle is proportional to the number of individuals sampled that share that haplotype. For mitochondrial haplotypes, shading denotes oceanic region as in Figure 3-1: Indian Ocean (black fill), western/central Pacific Ocean (open circles), and eastern Pacific Ocean (gray fill). Localities are listed for each nuclear haplotype as numbered in Figure 3-1 and Tables 3-1 through 3-3, increasing from west to east. For mitochondrial haplotype networks, Groups A-E show regions for which hypotheses of panmixia could be rejected by NCPA in favor of either restricted gene flow or contiguous range expansion (Table 3-4).

Figure 3-3. Mismatch distributions for mitochondrial haplotypes (1228 base pairs, black line) superimposed on the expected distribution calculated for the assumption of a demographically expanding population. Observed distributions do not differ significantly from expectations either in magnitude or raggedness (see Results). Peak values of approximately 17 (*G. flavimarginatus*) and 10 (*G. undulatus*) differences between haplotypes are consistent with an interpretation that some geographically isolated haplotype lineages formed during transitory geographic fragmentation of populations approximately 0.3 to 0.6 MY ago and have persisted through subsequent merging of populations.
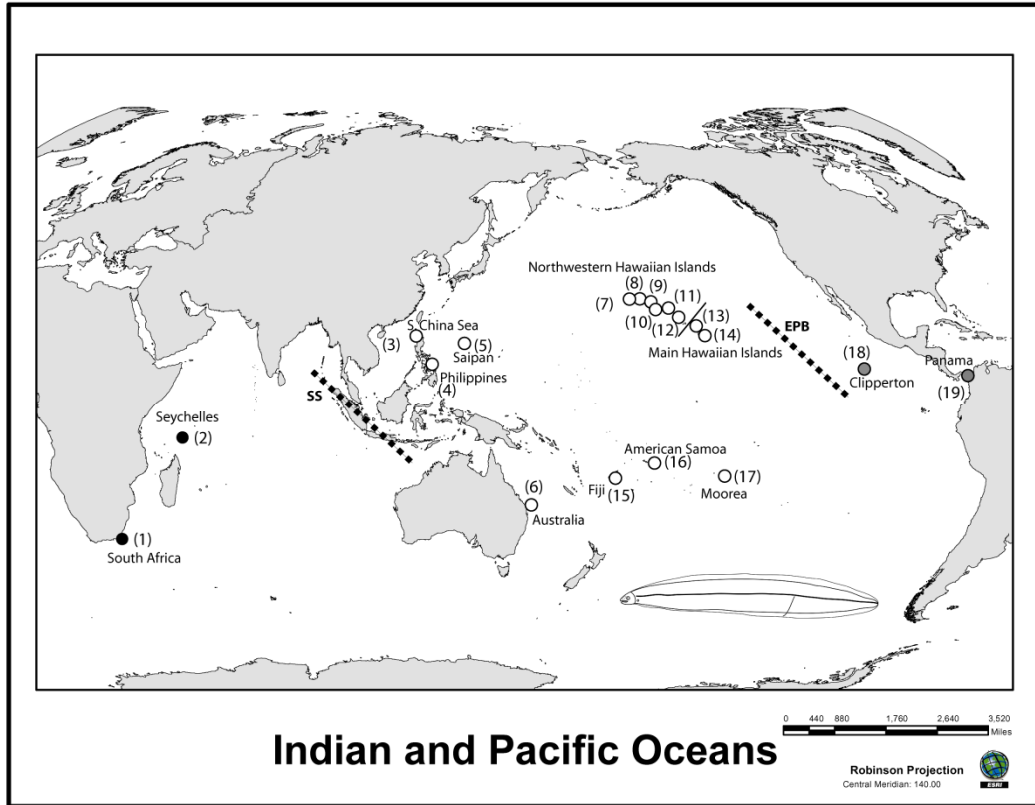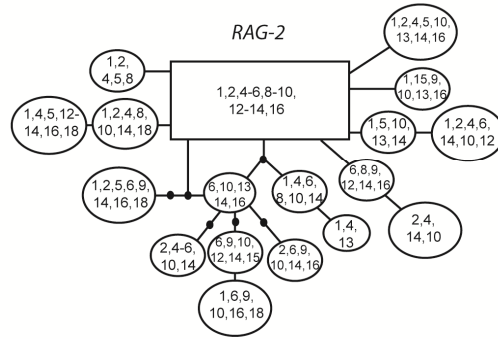
Indian and Pacific Oceans
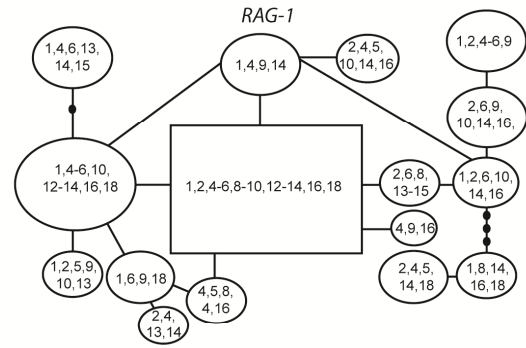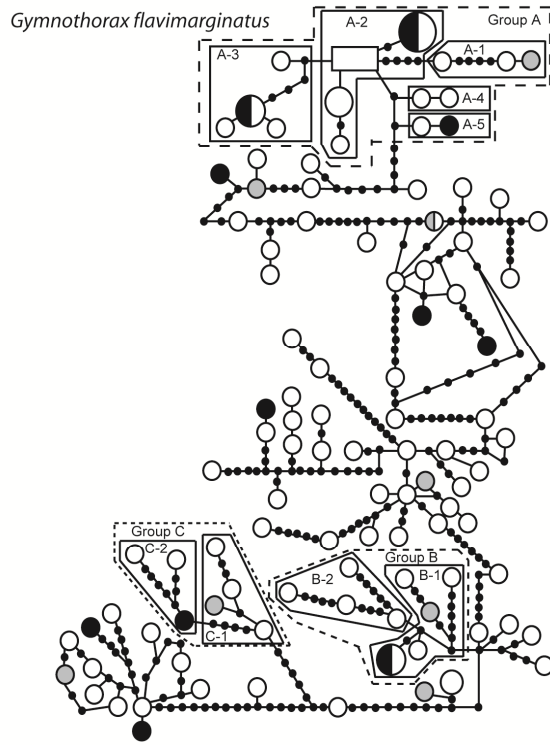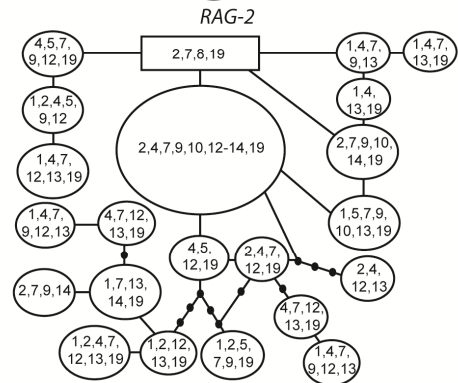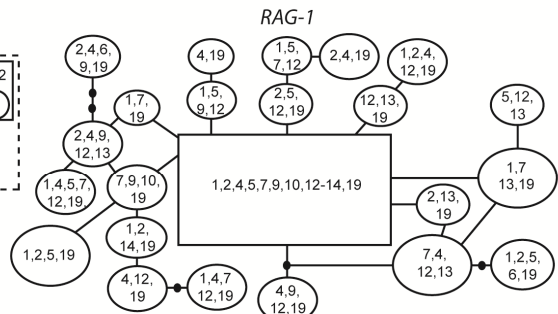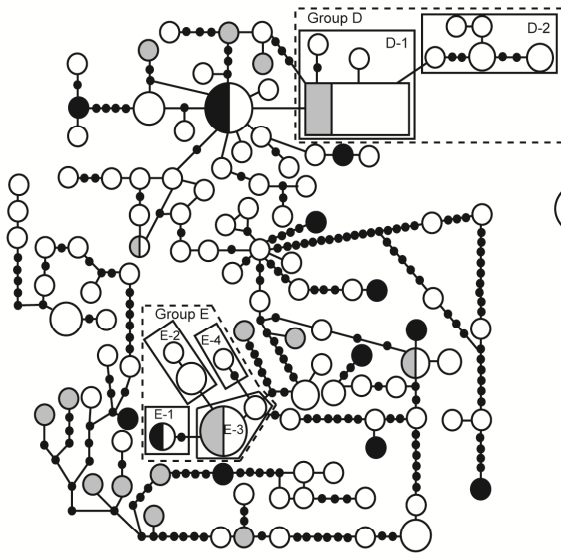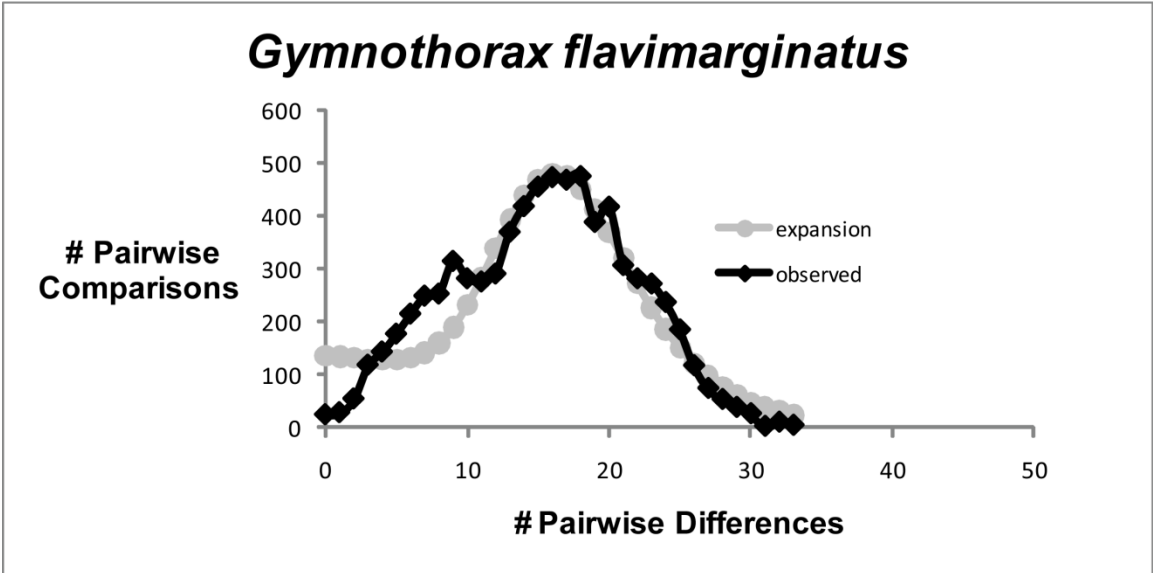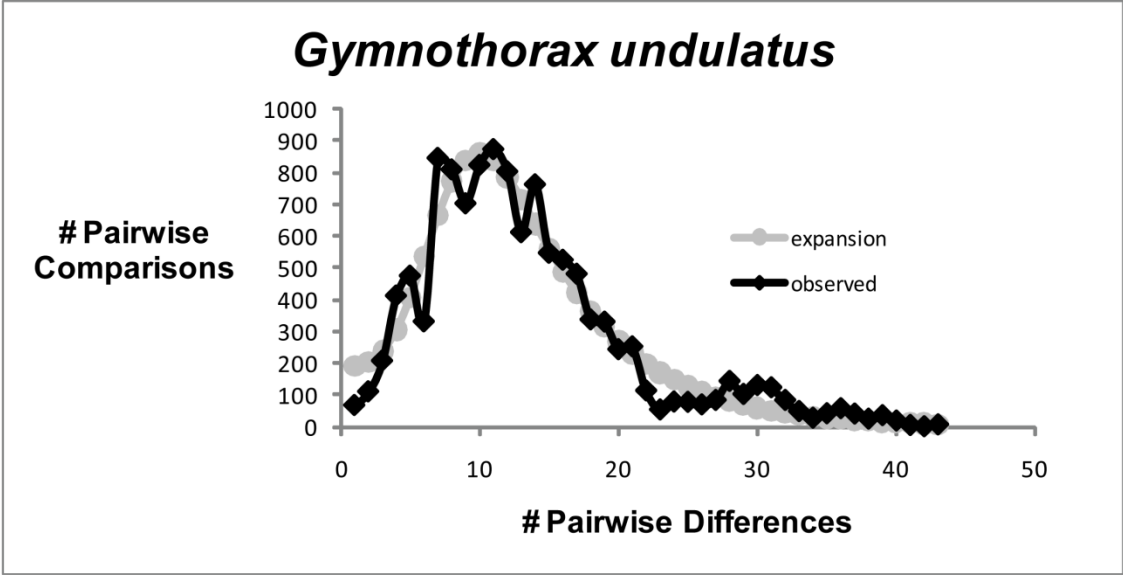
Figure 3-1

Figure 3-2

Figure 3-3

Literature Cited

Atema, J., Kingsford, M.J., Gerlack, G., 2002. Larval reef fish could use odour for detection, retention and orientation to reefs. Marine Ecology Progress Series 241, 151-160.

Barber, P.H., Erdmann, M.V., Palumbi, S.R., 2006. Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the coral triangle. Evolution 60, 1825-1839.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistics Society Proceedings B 57, 289-300.

Bermingham, E., McCafferty, S.S., Martin, A.P., 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian isthmus. In: Kocher, T.D., Stepien, C.A. (Eds.), Molecular systematics of fishes. New York: Academic Press; p. 113-128.

Bernardi, G., 2000. Barriers to gene flow in *Embiotoca jacksoni*, a marine fish lacking a pelagic larval stage. Evolution 54, 226-237.

Bernardi, G., Holbrook, S.J., Schmitt, R.J., Crane, N.L., 2003. Genetic evidence for two distinct clades in the French Polynesian population of coral reef three-spot damselfish *Dascyllus trimaculatus*. Marine Biology 143, 485-490.

Böhlke, E.B., McCosker, J.E., Böhlke, J.E., 1989. Family Muraenidae. In: Böhlke, E.B. (Ed.), Fishes of the Western North Atlantic. New Haven: Sears Foundation for Marine Research.

Bonferroni, C.E., 1936. Teoria statistica delle classi e calcolo delle probability a`
Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di
Firenze 8, 3-62.

Bowen, B.W., Bass, A.L., Muss, A., Carlin, J., Robertson, D.R., 2006a. Phylogeography
of two Atlantic squirrelfishes (Family Holocentridae): exploring links between
pelagic larval duration and population connectivity. Marine Biology 149, 899-913.

Bowen, B.W., Bass, A.L., Rocha, L.A., Grant, W.S., Robertson, D.R., 2001.
Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global
scale. Evolution 55, 1029-1039.

Bowen, B.W., Karl, S.A,. 2007. Population genetics and phylogeography of sea turtles.
Molecular Ecology 16, 4886-4907.

Bowen, B.W., Muss, A., Rocha, L.A., Grant, W.S., 2006b. Shallow mtDNA coalescence
in Atlantic Pygmy Angelfishes (Genus *Centropyge*) indicates a recent invasion from
the Indian Ocean. Journal of Heredity 97, 1-12.

Briggs, J.C., 1961. The East Pacific Barrier and the distribution of marine shore fishes.
Evolution 15, 545-554.

Briggs, J.C,. 1999. Coincident biogeographic patterns: Indo-West Pacific Ocean.
Evolution 53, 326-335.

Castonguay, M., 1987. Growth of American and European eel leptocephali as revealed
by otolith microstructure. Canadian Journal Zoology 65, 875-878.

Castro, A.L.F., Steward, B.S., Wilson, S.G., Hueter, R.E., Meekan, M.G., Motta, P.J., Bowen, B.W., Karl, S.A., 2007. Population genetic structure of Earth's largest fish, the whale shark (*Rhincodon typus*). Molecular Ecology 16, 5183-5192.

Clement, M., Posada, D., Crandall, K., 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9, 1657-1660.

Crabtree, R.E., Cyr, E., Bishop, R.E., Falkenstein, L., Dean, J.M., 1992. Age and growth of tarpon, *Megalops atlanticus*, larvae in the eastern Gulf of Mexico, with notes on relative abundance and probable spawning areas. Environmental Biology Fishes 35, 361-370.

Craig, M.T., Eble, J.A., Bowen, B.W., Robertson, D.R., 2007. High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). Marine Ecology Progress Series 334, 245-254.

Crisp, D.J., 1978. Genetic consequences of different reproductive strategies in marine invertebrates. In: Battaglia, B., Beadmore, J.A. (Eds.), Marine organisms: genetics, ecology and evolution. New York: Plenum; p. 257-273.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7, 214-221.

Eble, J.A., Toonen, R.J., Bowen, B.W., 2009. Endemism and dispersal: comparative phylogeography of three surgeonfish species across the Hawaiian Archipelago. Marine Biology 156, 689-698.

Ekman, S., 1953. Zoogeography of the sea. London: Sidgwick & Jackson Ltd.

Excoffier, L., Schneider, S., 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47-50.

Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479-625.

Floeter, S., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W., Wirtz, P., Edwards, A.J., Barreiros, G., Ferriera, J.P., Gasparini, J.L., Brito, A., Falcon, M., Bowen, B.W., Bernardi, G., 2008. Atlantic reef fish biogeography and evolution. Journal of Biogeography 35, 22-47.

Fu, Y-X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915-925.

Gaither, M.R., Toonen, R.J., Robertson, D.R., Planes, S., Bowen, B.W., 2010. Genetic evaluation of marine biogeographic barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus* spp.). Journal of Biogeography 37, 133-147.

Garrick, R.C., Dyer, R.J., Beheregaray, L.B., Sunnucks, P., 2008. Babies and bathwater: a comment on the premature obituary for nested clade phylogeographic analysis. Molecular Ecology 17, 1401-1403.

Harpending, R.C., 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human Biology 66, 591-600.

Hellberg, M.E., 2007. Footprints on the water: the genetic wake of dispersal among reefs. Coral Reefs 26, 463-473.

Hiatt, R.W., Strasburg, D.W., 1960. Ecological relationships of the fish fauna on coral reefs of the Marshall Islands. Ecological Monographs 30, 65-127.

Horne, J.B., van Herwerde, L., Choat, H.J., Robertson, D.R., 2008. High population connectivity across the Indo-Pacific: congruent lack of phylogeographic structure in three reef fish congeners. Molecular Phylogenetics and Evolution 49, 629-638.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2003. Basal actinopterygian relationships: a mitogenomic perspective on the phylogeny of the "ancient fish". Molecular Phylogenetics and Evolution 26, 110-120.

Ishikawa, S., Aoyama, J., Tsukamoto, K., Nishida, M., 2001. Population structure of the Japanese eel *Anguilla japonica* examined by mitochondrial DNA sequencing. Fisheries Science 67, 246-253.

Jones, G.P., Planes, S., Thorrold, S.R., 2005. Coral reef fish larvae settle close to home. Current Biology 15, 1314-1318.

Kingsford, M.J., Leis, J.M., Shanks, A., Lindeman, K.C., Morgan, S.G., Pineda, J., 2002. Sensory environments, larval abilities and local self recruitment. Bulletin of Marine Science 70, 309-340.

Klanten, O.S., Choat, H.J., Van Herwerden, L., 2007. Extreme genetic diversity and temporal rather than spatial partitioning in a widely distributed coral reef fish. Marine Biology 150, 659-670.

Knowles, L.L., 2008. Why does a method that fails continue to be used? Evolution 62, 2713-2717.

Leis, J.M., 1991. The pelagic stage of reef fishes: the larval biology of coral reef fishes. In: Sale, P.F. (Ed.). The ecology of fishes on coral reefs. New York: Academic Press; p. 183-230.

Leis, J.M., 2007. Behavior of fish larvae as an essential input for modeling larval dispersal: behavior, biogeography, hydrodynamics, ontogeny, physiology and phylogeny meet hydrography. Marine Ecology Progress Series 347, 185-193.

Leis, J.M., Wright, K.J., Johnson, R.N., 2007. Behavior that influences dispersal and connectivity in the small, young larvae of a reef fish. Marine Biology 153, 103-117.

Lessios, H.A., Robertson, D.R., 2006. Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. Proceedings of the Royal Society B 273, 2201-2208.

Lester, S.E., Ruttenberg, B.I., 2005. The relationship between pelagic larval duration and range size in tropical reef fishes: a synthetic analysis. Proceedings of the Royal Society B 272, 585-591.

Misawa, K., Fujii, S., Yamazaki, T., Takahashi, A., Takasaki, J., Yanagisawa, M., Ohnishi, Y., Nakamura, Y., Kamatani, N., 2008. New correction algorithms for multiple comparisons in case-control multilocus association studies based on haplotypes and diplotype configurations. Journal of Human Genetics 53, 789-801.

Mochioka, N., Iwamizu, M., 1996. Diet of anguilloid larvae: leptocephali feed selectively on larvacean houses and fecal pellets. Marine Biology 125, 447-452.

Montgomery, J.C., Tolimieri, N., Haine, O.S., 2001. Active habitat selection by pre-settlement reef fishes. Fish and Fisheries 2, 261-277.

Muss, A., Robertson, D.R., Stepien, C.A., Wirtz, P., Bowen, B.W., 2001. Phylogeography of *Ophioblennius*: the role of ocean currents and geography in reef fish evolution. Evolution 55, 561-572.

Nei, M., 1987. Molecular Evolutionary Genetics. Oxford, New York: Columbia
University Press.

Paillard, D., 2006. What drives the ice age cycle? Science 313, 455-456.

Panchal, M., Beaumont, M.A., 2007. The automation and evaluation of Nested Clade
Phylogeographic Analysis. Evolution 61, 1466-1480.

Petit, R.J., 2008. The coup de grâce for nested clade phylogeographic analysis?
Molecular Ecology 17, 516-518.

Planes, S., Fauvelot, C., 2002. Isolation by distance and vicariance drive genetic
structure of a coral reef fish in the Pacific Ocean. Evolution 56, 378-399.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics:
advantages of Akaike Information Criterion and Bayesian approaches over
Likelihood Ratio Tests. Systematic Biology 53, 793-808.

Posada, D., Crandall, K.A., Templeton, A.R., 2000. Geodis: a program for the cladistic
nested analysis of geographical distribution of haplotypes. Molecular Ecology 9,
487-488.

Rambaut, A., Drummond, A.J., 2007. Tracer v.1.4, Available from
http://beast.bio.ed.ac.uk/Tracer

Ramon, M.L., Nelson, P.A., DeMartini, E., Walsh, W.J., Bernardi, G., 2008.
Phylogeography, historical demography, and the role of post-settlement ecology in
two Hawaiian damselfish species. Marine Biology 153, 1207-1217.

Randall, J.E., 1998. Zoogeography of shore fishes of the Indo-Pacific region. Zoological
Studies 37, 227-268.

Randall, J.E., 2007. Reef and shore fishes of the Hawaiian Islands. Honolulu: University of Hawaii Sea Grant College Program; p. 546.

Riordan, C., Hussain, M., McCann, J., 2004. Moray eel attack in the tropics: a case report and review of the literature. Wilderness and Environmental Medicine 15, 194-197.

Rivera, M.A.J., Kelley, C.D., Roderick, G.K., 2004. Subtle population genetic structure in the Hawaiian grouper, *Epinephelus quernus* (Serranidae) as revealed by mitochondrial DNA analyses. Biological Journal of the Linnean Society 81, 449-468.

Rocha, L.A., Bass, A.L., Robertson, D.R., Bowen, B.W., 2002. Adult habitat preferences, larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes (Teleostei: Acanthuridae). Molecular Ecology 11, 243-251.

Rocha, L.A., Craig, M.T., Bowen, B.W., 2007. Phylogeography and the conservation of coral reef fishes. Coral Reefs 26, 501-512.

Rocha, L.A., Robertson, D.R., Rocha, C.R., Van Tassell, J.L., Craig, M.T., Bowen, B.W., 2005. Recent invasion of the tropical Atlantic by an Indo-Pacific coral reef fish. Molecular Ecology 14, 3921-3928.

Rosenberg, M.S., 2001. PASSAGE. Pattern Analysis, Spatial Statistics, and Geographic Exegesis. Version 1.1. Department of Biology, Arizona State University, Tempe, AZ.

Scheltema, R.S., 1968. Dispersal of larvae by equatorial currents and its importance to the zoogeography of shoal-water tropical species. Nature 217, 1159-1162.

Schmidt, J., 1923. The breeding places of the eel. Philosophical Transactions of the
Royal Society of London Series B 211, 179-208.

Schneider, S., Roessli, D., Excoffier, L., 2000. Arlequin ver. 2.000: a software for
population genetics data analysis. Genetics and Biometry Laboratory, University of
Geneva, Switzerland.

Schultz, J.K., Pyle, R.L., DeMartini, E., Bowen, B.W., 2007. Genetic connectivity
among color morphs and Pacific archipelagos for the flame angelfish, *Centropyge
loriculus*. Marine Biology 151, 167-175.

Song, C.B., Near, T.J., Page, L.M., 1998. Phylogenetic relations among percid fishes as
inferred from mitochondrial cytochrome *b* DNA sequence data. Molecular
Phylogenetics and Evolution 10, 343-353.

Springer, V.G., 1982. Pacific Plate biogeography, with special reference to shorefishes.
Smithsonian Contributions to Zoology 1-182.

Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype
reconstruction from population genotype data. American Journal of Human Genetics
73, 1162-1169.

Stephens, M., Smith, N., Donnelly, P., 2001. A new statistical method for haplotype
reconstruction from population data. American Journal of Human Genetics 68, 978-
989.

Swearer, S.E., Shima, J.S., Hellberg, M.E., Thorrold, S.R., Jones, G.P., Robertson, D.R.,
Morgan, S.G., Selkoe, K.Q., Ruiz, G.M., Warner, R.R., 2002. Evidence of self-
recruitment in demersal marine populations. Bulletin of Marine Science 70, 251-272.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA
polymorphism. Genetics 123, 585-595.

Taberlet, P., Meyer, A., Bouvet, J., 1992. Unusually large mitochondrial variation in
populations of the blue tit, *Parus caeruleus*. Molecular Ecology 1, 27-36.

Taylor, M.S., Hellberg, M.E., 2003. Genetic evidence for local retention of pelagic
larvae in a Caribbean reef fish. Science 299, 107-109.

Taylor, M.S., Hellberg, M.E., 2005. Marine radiations at small geographic scales:
speciation in neotropical reef gobies (*Elacatinus*). Evolution 59, 374-385.

Templeton, A.R., 1998. Nested clade analysis of phylogeographic data: testing
hypotheses about gene flow and population history. Molecular Ecology 7, 381-397.

Templeton, A.R., 2008. Reply: Nested clade analysis: an extensively validated method
for strong phylogeographic inference. Molecular Ecology 17, 1877-1880.

Templeton, A.R., 2009. Statistical hypothesis testing in intraspecific phylogeography:
nested clade phylogeographical analysis vs. approximate Bayesian computation.
Molecular Ecology 18, 319-331.

Thacker, C.E., Thompson, A.R., Roje, D.M., Shaw, E.Y., 2007. New expansions in old
clades: population genetics and phylogeny in a *Gnatholepis* species (Teleostei:
Gobiodei) in the Pacific. Marine Biology 153, 375-385.

Theisen, T.C., Bowen, B.W., Lanier, W., Baldwin, J.D., 2008. High connectivity on a
global scale in the pelagic wahoo, *Acanthocybium solandri* (tuna family
Scombridae). Molecular Ecology 17, 4233-4247.

Tseng, W., 1990. Relationship between growth rate and age at recruitment of *Anguilla japonica* elvers in a Taiwan estuary as inferred from otolith increments. Marine Biology 107, 75-81.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D., 2005. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society of London, Series B 360, 1847-1857.

Weersing, K.A., Toonen, R.J., 2009. Population genetics, larval dispersal, and demographic connectivity in marine systems. Marine Ecology Progress series 393, 1-12.

CHAPTER 4

PAGES 73-103

COMPARATIVE PHYLOGEOGRAPHY OF FOUR INDO-PACIFIC MORAY EEL SPECIES: HIGH PELAGIC LARVAL DURATION ENSURES OCEAN-WIDE GENETIC HOMOGENEITY OF SPECIES DESPITE VARIATION IN ADULT NICHE BREADTH

Introduction

There are approximately 200 species of moray eels (Muraenidae) globally. They occur in every ocean basin, and they have the lowest level of endemism and some of the largest range sizes among all reef fishes (Hourigan and Reese, 1987). Many species have ranges that include the entire east-west expanse of the Indo-Pacific (2/3rds of the planet), and within that range adult habitat can vary from very shallow water to reef habitats 600 meters deep (Böhlke et al., 1989; Böhlke and Smith, 2002). There are 150 species distributed in the Indo-Pacific, and many of them have fully sympatric distributions (Böhlke and Randall, 2000; Böhlke and Smith, 2002). Like many reef fishes, adult moray eels do not relocate over large distances (Böhlke et al., 1989), and all dispersal occurs by pelagic larvae. All morays have a larval form that confers enormous dispersal potential, relative to other reef fishes, through extended pelagic larval duration (PLD) (Thresher, 1984), the period during which most dispersal is accomplished in reef fishes.

Of species with large, sympatric distributions in the Indo-Pacific, some occupy nested subsets of the habitat types used by others. For example, *Gymnothorax undulatus* and *Gymnothorax flavimarginatus* occur from the west coast of the Americas to the east coast of Africa, and throughout that range adults occupy reefs and rocky substrates from depths of zero to 150 meters. Both of these species show high levels of gene flow throughout their ranges, across what are major biogeographic barriers to most species of reef fishes (Reece et al., 2010; Chapter 3). *Echidna nebulosa* and *Gymnomuraena zebra* occur throughout the Indo-Pacific, sympatric with the two *Gymnothorax* species, but within that range adults primarily occupy reefs between zero and 15m deep. The two *Gymnothorax* species are generalist predators that consume fish, cephalopods and

74

crustaceans (Harmelin-Vivien, 1979; Yukihira et al., 1994), whereas *Echidna nebulosa* and *Gymnomuraena zebra* both specialize on crustaceans, a subset of the *Gymnothorax* diet (Hiatt and Strasburg, 1960; Lieske and Myers, 1996; Randall, 1967; Yukihira et al., 1994). The depth of reef habitat used by a species (Chatfield et al., 2010; Fasola et al., 1997; Mehta, 2008; Sampayo et al., 2007) and diet (Gaughan and Potter, 1997; Scharf et al., 2000) are suitable proxies for niche breadth; for both proxies, the two *Gymnothorax* species are similar to each other in measured niche breadth, and *E. nebulosa* and *G. zebra* are nearly equivalent in occupying a nested subset of the niche dimensions of the *Gymnothorax* species. The codistributed species have subpopulations at the same geographic locations and share the same mechanism of dispersal among subpopulations by leptocephalus larvae, but the numbers of adult individuals and their local geographic distributions differ greatly among species. This geographic pattern permits controlled comparisons of the effect of niche breadth on population genetic structure in highly dispersive organisms.

Numerous studies have consistently identified niche breadth as a stronger determinant of range size and abundance than dispersal capability (Bonte et al., 2004; Brändle and Brandl, 2001; Hawkins et al., 2006; Heino and Soininen, 2006; Köckemann et al., 2009; Pino et al., 2008; Swihart et al., 2003; Swihart et al., 2006; Thompson et al., 1999). I am unaware of any study that measures the impact of niche breadth on population genetic structure within species using comparisons of species that have geographically congruent patterns of fragmentation into subpopulations but which differ in niche breadth. Habitat specialization by a species typically imposes higher levels of habitat fragmentation and reduced gene flow among subpopulations (Riginos and

Nachman, 2001; Soons et al., 2005; Young et al., 1996). Does habitat specialization by adult moray eels increase geographic genetic structuring within moray-eel species, or does their larval-dispersal mechanism erase any consequences of habitat constriction on geographic genetic structuring?

My prior work (Reece et al., 2010; Chapter 3) shows that two codistributed habitat generalists, *Gymnothorax undulatus* and *Gymnothorax flavimarginatus*, both maintain high levels of gene flow throughout their distributions, with some remnants of past geographic fragmentation evident in their haplotype variation. Estimated coalescence times for their mitochondrial genomic haplotypes exceed 1.7 million years for *G. flavimarginatus* and 3.5 million years for *G. undulatus.* Here, I present comparative phylogeographic analyses using orthologous mitochondrial and nuclear genomic markers for two species of habitat specialists (*Echidna nebulosa* and *Gymnomuraena zebra*). I quantify differences among these four species in expected spatial dimensions of their subpopulations using data from available collection records to estimate their depth tolerances, and geographical information systems to estimate their current geographic distributions. Although exact dimensions of adult subpopulations and relative areas occupied undoubtedly changed with global climatic cycles (see discussion by Reece et al., 2010; Chapter 3), I assume that the *Gymnothorax* species consistently occupied larger portions of the reefs than did *E. nebulosa* and *G. zebra*. Molecular phylogenetic analyses indicate that the four species compared in this study have evolved separately from each other for at least 20 million years, with the estimated separation of the *Gymnothorax* species from each other being 24.3 million years, and the estimated separation of *Gymnomuraena zebra* from a clade comprising the other three species being 27.2 million

years (Reece et al., in press; Chapter 2). Because the phylogenetic separation of these species greatly exceeds estimates of coalescence times of haplotype variation within them, my comparisons of geographic genetic structuring within these species should be free of any phylogenetic artifacts.

Methods

I estimated habitat utilization by depth for each species by examining records for museum specimens, field guides, and my own extensive field sampling. The online resource Fishbase.org, a clearing-house for published species records and museum collections worldwide, was an important source for specimen locality and depth data. Most museum vouchers of moray eels are small specimens collected at depths shallow enough to employ rotenone and dipnets, so museum specimens are biased towards shallow sampling sites. I recorded the mean and maximum depths sampled from Fishbase.org and confirmed depths directly from specimen tags in the Smithsonian's ichthyology collection in Washington, D.C. I translated the depths of reefs occupied by a species into a quantitative measure of niche (habitat) breadth using ArcGIS v.9 and bathymetric data (database ETOPO1) from the National Oceanic and Atmospheric Association National Geophysical Data Center. Resolution of these data is 1 arc-minute global relief, depicting topography of the Earth's surface, including ocean bathymetry (http://www.ngdc.noaa.gov). ETOPO1 data were imported into ArcGIS and rasterized (translates pixels of an image into a grid of depth information). Topographic data were binned to exclude all elevations above sea level and deeper than 150 meters. I binned the remaining data into three elevation classes: zero to -1m, -1m to -15m, and -15m to -150m. I included areas between zero and -1m for graphical representation of coastlines,

but these areas were excluded from my calculations to avoid counting tidal areas. Range data for each species were taken from Fishbase.org point data, which include latitude and longitude for every museum specimen recorded for each species. All four species had largely overlapping distributions; these distributions were combined into a single range, which was then manually inscribed within a polygon. Within the polygon representing the ranges of all four species, I computed the area in square meters for each of the three depth classes defined above. All areas between -1 and -15 meters altitude were included for *E. nebulosa* and *Gymnomuraena zebra*. All areas between -1 and -150 meters were calculated for the two *Gymnothorax* species. I express the portion of habitat used by the habitat specialists as a percentage of the total habitat used by the generalist *Gymnothorax* species.

One key assumption of this study is that dispersal capabilities do not differ substantially among the four species of moray eels, because the effects of differences in niche breadth might be confounded or mitigated by differences in dispersal capability. Dispersal capabilities for Muraenidae were inferred from the literature when available. I examined otoliths of each of the remaining three species. Otolith bones contain ringed deposits that can be counted to measure age preceding metamorphosis from the larval to the adult stage, thereby quantifying the amount of time spent in the pelagic larval stage, called the pelagic larval duration (see Kuroki et al., 2006).

Tissue samples were collected through a combination of spearfishing, baited lobster traps, museum collections, and local fishermen. When possible, each sample collected for this study included data on the depth from which the animal was sampled. DNA isolation and sequence generation were performed exactly according to previously

published methods (Reece et al., 2010; Chapter 3). Haplotype networks were reconstructed in TCS v.1.2.1 (Clement et al., 2000). Tajima's D test for neutrality was performed in Arlequin v.3.1.1 (Excoffier and Schneider, 2005). To test for phylogeographic structure in *Echidna nebulosa* and *Gymnomuraena zebra*, pairwise comparisons of genetic distances among species, pairwise Φ_{ST} comparisons within species, and AMOVA within species were completed in Arlequin. For AMOVA analyses, I divided individuals into geographic localities, biogeographic regions separated by the Sunda Shelf and eastern Pacific Barrier, and all populations combined. Pairwise Φ_{ST} comparisons included all populations with more than five samples for mtDNA, *RAG-1*, and *RAG-2*, with a B-H (Benjamini and Hochberg, 1995) correction for multiple comparisons. I tested for isolation by distance through Mantel tests comparing a matrix of pairwise Φ_{ST} values with a matrix of linear geographic distances between sites in the program Passage (Rosenberg, 2001). I inferred the history of demographic expansion/contraction through Fu's F_S (Fu, 1997), which tests the null hypothesis of constant population size, and mismatch distributions of mtDNA, which tests the null hypothesis of demographic expansion, at an alpha of 0.05.

I estimated coalescence times for genetic variation within *E. nebulosa* and *G. zebra* separately in the program BEAST v.1.5.4 (Drummond and Rambaut, 2007). Estimates for *G. undulatus* and *G. flavimarginatus* were calculated through the same methods and input from a previous study (Reece et al., 2010; Chapter 3). A model of evolution for each codon position of each of the four markers was computed in jModelTest (Posada, 2008) by the AIC method. Mutation rates were taken for each region and are based on fossil and biogeographic calibration points (Reece et al., in press;

Chapter 2). I used 30 million generations with a sampling interval of 2000. Two independent runs for each species were determined to have reached stationarity when ESS values were above 200, as visualized in the program Tracer v.1.5.0 (Rambaut and Drummond, 2007). The runs were combined using LogCombiner v.1.5.4 (Drummond and Rambaut, 2007), and the trees were created in TreeAnnotatotor 1.5.4 (Drummond and Rambaut, 2007) and visualized in FigTree 1.3.1 (http://tree.bio.ed.ac.uk/). I report on coalescence times for each species in the form of 95% posterior probability densities (PPD) surrounding those estimates.

Because moray eels are notoriously difficult to sample, my sample sizes are low by most standards. However, my sampling permits sufficient statistical power to reject for all pairwise population comparisons the hypothesis of a migration rate low enough to produce isolation by distance. I performed simulations of gene flow among populations for *E. nebulosa* and *G. zebra* in the program Lamarc 2.1.3 (Kuhner, 2006). For each species, I conducted pairwise estimates of gene flow between: 1) each geographic population, and 2) between each biogeographic region as delimited by the Sunda Shelf and the eastern Pacific Barrier (Reece et al., 2010; Chapter 3). I simulated gene genealogies based on observed patterns of DNA variation within each of the narrow-niche species, and modeled the levels of gene flow among populations that would yield the patterns I observed in my data. Because Lamarc estimates directional migration, the numbers of migrants for each direction were summed to estimate total bidirectional migration between each of the two populations. Parameters for the Lamarc file were as follows: effective population sizes were scaled to 0.25 for mtDNA and 1 for nuclear markers; a GTR model was used for all partitions; relative mutation rates were input per

previous findings (Reece et al., 2010; Chapter 3); I used a Bayesian analysis of four

chains with heating (induces exploration of parameter space following occupation of a

local maximum); two replicates, an initial chain of 1000 with sampling interval of 20 and

a final chain of five million with a sampling interval of 200 and a burn-in of 1000. The

goal of this analysis was to yield simulation-based estimates of gene flow among

populations sufficient to prevent isolation by distance, which I consider an Nm greater

than 5 per generation, where N is the variance effective size of the population and m is

the proportion of each population exchanged as migrants to the other one. This is a

conservative criterion, as an Nm of 1 is typically considered a sufficient level of gene

flow to prevent isolation by distance (Templeton, 2006). The parameter Nm is estimated

per Lamarc documentation as the parameter theta (two times the mutation rate per

generation times the number of copies of homologous DNA in the gene pool, which is

4Nμ) multiplied by the parameter M (= m/μ, migration rate, or the chance for a gene

lineage to immigrate per generation divided by the mutation rate per site per generation),

which yields 4Nm, or the number of diploid immigrants per generation.

Results

Depth Range and Dispersal Capabilities for Each Species

    Gymnothorax flavimarginatus is reportedly captured consistently between 1 and

150 meters below sea level (Myers, 1991). The maximum depth reported on Fishbase.org

is 150m, and the average is 10.5m. *Gymnothorax undulatus* is consistently observed

between 1 and 150m depth (Mundy, 2005; Randall, 2007). The deepest record on

Fishbase.org is 150m, and the average depth is 10.5m. Most *Gymnomuraena zebra*

specimens were observed between zero and 15m (Mundy, 2005; Sommer et al., 1996).

The average depth reported on Fishbase.org and confirmed through the Smithsonian

collection is 1.2m, and the maximum depth reported was 10.6m. The reported depth

range for *Echidna nebulosa* is also between 1 and 15m (Kuiter, 1998; Mundy, 2005), and

the average depth reported on Fishbase.org and in the Smithsonian collection is 2.0m,

with a maximum depth recorded at 15.2m. There are reliable accounts of *E. nebulosa* and

G. zebra being caught at depths up to 40m (e.g., Randall, 2007); however, each of these

species primarily inhabit shallow waters and upwards of 90% of specimens in major

museums with precise collection records were captured at depths shallower than 5m. I

collected 79 specimens of *E. nebulosa* through my own efforts, through collaborators,

and from the Pet Trade. The average depth of *E. nebulosa* caught was 1.8m. I collected

67 specimens of *G. zebra*, with an average depth recorded of 1.2m.

Within the range of all four species (Figure 4-1), ArcGIS analyses of the

bathymetric landscape estimated 650,000 m^2 of suitable habitat by depth between -1 and -

15m. I estimated 2.6 million m^2 of habitat between -1 and -150m. A simple calculation

demonstrates that *E. nebulosa* and *Gymnomuraena zebra* could at most occupy

approximately 20% of the habitat available to *Gymnothorax undulatus* and *Gymnothorax

flavimarginatus*.

Estimates of pelagic larval duration (PLD), a proxy for dispersal capability

(Lester and Ruttenberg, 2005), are available for *E. nebulosa* at greater than 80 days

(Brothers and Thresher, 1985). My attempts to estimate PLDs from the otoliths of the

other three species merely confirmed reports that muraenid otoliths are extremely

difficult to interpret, but my results clearly indicate long pre-metamorphic durations

typically greater than 80 days. This value is consistent with others reported across

multiple genera of morays (Bishop et al., 2000; Böhlke et al., 1989; Brothers and

Thresher, 1985; Leis, 1991; Leis and Miller, 1976; Thresher, 1984). While difficult to

estimate precisely, based on my examination of otoliths and those reported in the

literature, the PLDs of the four species in question are all well beyond the 15-30 day

average for most reef fishes (Lester and Ruttenberg, 2005).

DNA Sequence Data for Echidna nebulosa and Gymnomuraena zebra

I sampled portions of two mitochondrial genes (632 bp of *CYB* and 596 bp of

COI; treated as a single locus) and two nuclear genes (420 bp of *RAG-1* and 746 bp of

RAG-2) for a total of 2,394 base pairs of DNA. Haplotype networks are shown in Figure

4-2. Table 4-1 shows the numbers of alleles for each marker and each species. Tajima's

D test fails to reject neutrality for all markers in both species (all *P*-values > 0.1).

The demographic histories of both species include evidence for population

expansion (Table 4-1). Fu's F_S values are significantly negative (providing evidence for

demographic expansion), and observed mismatch distributions resemble simulated

distributions under a model of demographic expansion (Table 4-1). Estimates of

coalescence times from analyses in BEAST (Table 4-1) demonstrate that geographic

patterns of genetic diversity in both species evolved within the last 2.2 to 9.1 MY (95%

PPD). These estimates are broadly overlapping with those reported for the two

Gymnothorax species (Table 4-1). Pairwise Φ_{st} comparisons for mtDNA and both nuclear

loci failed to identify significant comparisons, indicating minimal population structure as

a result of high gene flow throughout the species' ranges (mtDNA results presented in

Table 4-2). Analyses of Molecular Variance (AMOVA) demonstrate that over 95% of genetic variation in each of the four species occurs within populations, indicating little to no geographic population structure. The amount of total genetic variation explained by oceanic regions delimited by the Sunda Shelf and Eastern Pacific Barrier (Φ_{SC}) is 0.3% and 0.1% of the variation for *G. zebra* and *E. nebulosa*, respectively, and not significantly different from zero. Tests for isolation by distance found no significant association between geographic and genetic distances among populations based on a Mantel test (all $R < 0.001$, *P*-values > 0.3).

I further investigated evidence for population structure by modeling gene flow under coalescent simulations in the program Lamarc v.2.1.3 (Kuhner, 2006). For *Echidna nebulosa* and *Gymnomuraena zebra*, I statistically reject the hypothesis that any pair of populations exchanges on average fewer than five migrants in each direction per generation. At the broadest scale, estimates of migration between regions delimited by the biogeographic barriers of the Sunda Shelf and the Eastern Pacific Barrier are approximately 10 migrants per generation between each region. I present the posterior probability densities of migration estimates between biogeographic regions for *Echidna nebulosa* (Figure 4-3), with similar results of greater than 10 migrants per generation for *G. zebra* (not shown). These levels of gene flow clearly support the traditional metrics of population structure that indicate low levels of structure and high gene flow.

Discussion

Although the shallow-water moray species *Echidna nebulosa* and *Gymnomuraena zebra* occupy only 20% of the reef areas occupied by habitat generalists *Gymnothorax*

flavimarginatus and *Gymnothorax undulatus*, they maintain comparably high levels of gene flow on the same ocean-wide geographic scale reported for the *Gymnothorax* species (Reece et al., 2010; Chapter 3). Haplotypes from the Indian, western/central Pacific, and eastern Pacific Oceans are interspersed on the networks for all three loci (mtDNA, *RAG-1, RAG-2*) with almost no geographic clustering of haplotypes (Figure 4-2). Geographic genetic homogeneity seems even more extreme for the two shallow-water species than for the habitat generalists. As summarized by Φ_{SC}, the total proportion of genetic variation that exists as differences among populations is only 0.3% for *G. zebra* and only 0.1% for *E. nebulosa* despite high levels of variation at all loci (Table 4-1). Minimum coalescence times estimated for genetic variation within *E. nebulosa* (4.2 Myr) and *G. zebra* (2.2 Myr) span multiple climatic cycles in which the Sunda Shelf would have emerged as a land barrier separating the Indian Ocean from the Pacific Ocean (see discussion by Reece et al., 2010; Chapter 3). My molecular genetic analysis of the *Gymnothorax* species showed some remnants of past geographic fragmentation of the oceans in some young haplotype clades that had not yet reached ocean-wide distribution (Reece et al., 2010; Chapter 3), but remnants of past fragmentation are much less evident in the results presented here.

Haplotype networks for *Gymnomuraena zebra* (Figure 4-2) show almost no signs of any past fragmentation of populations and provide the most extreme case of geographically widespread genetic homogeneity yet reported for any reef fish species dependent on larval dispersal. Of the 64 mitochondrial haplotypes sampled, the inferred ancestral haplotype is sampled one time each in the Indian Ocean and the central Pacific, and all other haplotypes are sampled from individual specimens only. There is no

tendency for mitochondrial haplotypes to cluster with others from the same oceanic region. A high level of mutation for the mitochondrial genome combined with evidence for a demographically expanding population (Table 4-1) explains the result that almost every specimen sampled had a different mitochondrial haplotype. For the more slowly evolving nuclear markers, the most common haplotypes typically occur in all ocean regions. Of the 47 haplotypes sampled collectively by the two nuclear loci, twenty haplotypes span all three oceanic regions and only a single rare haplotype is restricted to a single region, the central Pacific. Although the mitochondrial and nuclear genomic haplotype networks represent opposite extremes in sharing of haplotypes among the three oceanic regions, they convey a common message of widespread gene flow across geographic populations in a species undergoing demographic expansion.

Haplotype networks for *E. nebulosa* show some traces of past geographic fragmentation of populations, although less so than the corresponding networks for the *Gymnothorax* species (Reece et al., 2010; Chapter 3). Of the 67 mitochondrial haplotypes scored for this species, the inferred ancestral haplotype and several groups of haplotypes at various tips of the network are from the central Pacific. A grouping of 12 haplotypes that form a clade shown at the bottom of the network (Figure 4-2) includes some deeply divergent haplotypes and lacks members from the Indian Ocean. These patterns plus the occurrence of central Pacific haplotypes throughout the network indicate that the central and eastern Pacific have been a reservoir of genetic diversity for *E. nebulosa.* Of the 44 haplotypes sampled collectively for the two nuclear markers, four rare haplotypes are sampled from the central Pacific only, and one haplotype is sampled only from the Indian Ocean. The prevailing pattern for the nuclear markers is nonetheless one of sharing of

haplotypes among the oceanic regions. Nineteen of the 44 nuclear haplotypes sampled from *E. nebulosa* occur among the three oceanic regions, representing localities separated by more than 22,000 kilometers and two major biogeographic barriers.

These data confirm predictions of my prior study (Reece et al. 2010; Chapter 3) that the leptocephalus larva of moray eels permits ocean-wide exchange of genes among populations. The form of the leptocephalus is likely an adaptation for permitting morays to find sparsely distributed habitats suitable for adults. The effectiveness of this mechanism is evident in the finding that the greater habitat fragmentation characteristic of *E. nebulosa* and *G. zebra* adults does not decrease maintenance of high levels of gene flow throughout the Pacific or the maintenance of high levels of genetic variation for both mitochondrial and nuclear genomes. Estimated coalescence times indicate that current haplotype variation represents at least several million years of accumulation. If the smaller geographic areas of the *E. nebulosa* and *G. zebra* adult populations had produced greater local extinction and recolonization of populations relative to the *Gymnothorax* species, *E. nebulosa* and *G. zebra* should show lower levels of variation and later coalescence times than the *Gymnothorax* species, and they do not. All four moray species show very similar levels of nuclear haplotypic variation (Table 4-1). For all species, most individuals sampled have mitochondrial haplotypes not found in any other individuals; the number of mitochondrial haplotypes reported therefore closely tracks sample size (Table 4-1), with no indication that the species differ greatly in mitochondrial genomic variation. The wide dispersal of leptocephalus larvae permits moray species to maintain similarly high levels of genetic variation despite a five-fold difference in the amount of suitable habitat available to adult individuals.

Following Horne et al. (2008), Klanten et al. (2007), and Reece et al. (2010; Chapter 3) I use the term "temporal genetic partitioning" to denote a genetic pattern in Indo-Pacific morays of population genetic homogeneity on a million-year time scale with transitory fragmentations and demographic expansions on shorter time scales. Global climatic fluctuations are the most obvious source of recurring cycles of fragmentation followed by expansion of moray populations. Randall (1998) hypothesizes three to six episodes within the past 700,000 years during which sea level would have been low enough to expose the Sunda Shelf and thereby to preclude dispersal of leptocephali between the Indian and Pacific Oceans. Climatic cycles should have less impact on Eastern Pacific Barrier than they do on the Sunda Shelf unless changing sea levels greatly disrupt currents carrying leptocephali between the eastern and central Pacific. Geological climatic data reviewed by Paillard (2006) indicate sea-level fluctuations occurring on a cycle of ~100,000 years throughout the past million years. Reece et al. (2010; Chapter 3) interpreted geographic restriction of some younger mitochondrial haplotypes in the two *Gymnothorax* species as evidence that they are emerging from a transitory fragmentation event that occurred over the past 0.3 to 0.6 MY. Geographic restriction of some younger haplotype clades of *E. nebulosa* to the central pacific and the central/eastern Pacific likewise can be attributed to temporal genetic partitioning. Although the same transitory barriers would have fragmented *G. zebra* across the Sunda Shelf, subsequent genetic mixture has obliterated all traces of past fragmentation in this species, although its comparably high levels of haplotype variation and deep coalescence times indicate that it shares the temporal geographic history proposed for the other morays.

My results predict occurrence of widespread geographic genetic homogeneity in all moray species that have leptocephali with comparably long pelagic durations. The pelagic larva is of prime importance in explaining the history of variation in morays as in many other marine fishes (but see Weersing and Toonen, 2009) and in this case it is sufficiently powerful to supersede factors, such as adult niche breadth, that otherwise would be extremely limiting to gene flow. The climatic cycles that occur on an approximately 100,000 year timescale produce transitory geographic fragmentation and differentiation of populations across the Sunda Shelf, usually followed by genetic homogenization of species across the Indo-Pacific when the barrier is removed. These periods of geographic genetic fragmentation nonetheless provide the most parsimonious hypothesis for speciation and lineage accumulation in moray eels. Should reproductive isolation occur at a low rate between geographically isolated populations, subsequent inundation of the Sunda Shelf would permit leptocephali of both daughter species subsequently to acquire ocean-wide distributions. Accumulation of moray species lineages with ocean-wide distributions thus occurs on a timescale of millions of years.

Table 4-1. Number of individuals sampled, the number of alleles (mtDNA, *RAG-1*, *RAG-2*), coalescence time for standing variation (TMRCA = time to most recent common ancestry) based on all 3 gene regions, portion of genetic variance among biogeographic regions (Φ_{SC}) based on mtDNA, and evidence of demographic expansion in the form of Fu's F_S (significant Φ_{SC} and F_S values in bold) for mtDNA. Mitochondrial DNA was the only region variable enough to be informative for AMOVA, Fu's F_S and mismatch distributions. A significant negative F_S indicates that the null hypothesis of constant population size is rejected in favor of demographic expansion, while non-signficant *P*-values from mismatch distributions are consistent with a null hypothesis of demographic expansion. No values of Harpending's Raggedness Index (HRI) are significant at α=0.05, demonstrating that the data do not differ from simulated results in their fit to models of population growth. Data for *G. undulatus* and *G. flavimarginatus* are from Reece et al. (2010; Chapter 3).

Species	N	# alleles	TMRCA	Φ_{SC}	Fu's F_S	Mismatch Distributions
Gymnothorax undulatus	165	112,23,22	3.5-8.6 MY	**0.039**	**-24**	*P* = 0.42 HRI=0.20
Gymnothorax flavimarginatus	124	104,16,18	1.7-3.0 MY	0.009	**-24**	*P* = 0.78 HRI=0.10
Gymnomuraena zebra	67	64, 23,24	2.2-5.3 MY	0.003	**-57**	*P* = 0.77 HRI=0.002
Echidna nebulosa	79	67,19,25	4.2-9.1 MY	0.001	**-45**	*P* = 0.68 HRI=0.002

Table 4-2. Pairwise Φ_{st} values for *Gymnomuraena zebra* (below diagonal) and *Echidna nebulosa* (above diagonal) for mtDNA. None of the comparisons were significant with or without the B-H correction for multiple comparisons, indicating no population structure. No significant pairwise comparisons were observed in either of the nuclear markers for either species (data not shown). Negative values have no biological interpretation are statistically indistinguishable from zero. The two columns on the right represent sample sizes for *E. nebulosa* (Ene) and *G. zebra* (Gze), with the exception of 1 individual *E. nebulosa* sampled from Moorea.

	Clipperton	Panama	Hawaii	Oahu	Australia	Saipan	Indonesia	Philippines	Taiwan	S. Africa	Ene	Gze
Clipperton		-0.39	-0.50	-0.41	-0.43	-0.58	-0.49	-0.31	-0.38	-0.45	1	4
Panama	-0.03		-0.03	-0.02	0.00	-0.05	-0.05	-0.01	-0.04	-0.02	15	14
Hawaii	0.01	-0.06		-0.03	-0.10	-0.13	-0.05	0.02	-0.03	-0.05	9	10
Oahu					-0.03	-0.06	-0.06	-0.06	-0.07	-0.04	6	0
Australia						-0.16	-0.01	0.04	0.00	-0.06	5	0
Saipan							-0.10	-0.01	-0.06	-0.09	5	0
Indonesia	-0.16	0.01	0.03					-0.02	-0.08	-0.05	7	5
Philippines	-0.07	-0.01	0.00				-0.04		0.01	-0.01	10	13
Taiwan							-0.08			-0.07	5	2
S. Africa	-0.05	-0.05	-0.04				-0.02	-0.04			15	14
Seychelles	0.04	0.02	-0.02				0.01	0.07		0.03	0	5
										Totals	79	67

Figure Legends

Figure 4-1. Map of sampling sites and potential habitat by depth. Habitat depths are given in grey and black and are limited to the geographic range reported for the four species (see methods). All land above sea level and bathymetry data below -200m has been removed with the exception of coastline. Two insets depict the the habitat range by depth contrasting for the two broad-niche species (*Gymnothorax spp.*) in grey+black, and the two narrow-niche species (*Echidna nebulosa* and *Gymnomuraena zebra*) in black only. Note that in both ocean islands (top left) and along continental shelfs (lower right), the shallow habitat (1-15m) is a dramatically reduced subset of the shallow+deep habitat (15-150m).

Figure 4-2. Haplotype networks for *Echidna nebulosa* and *Gymnomuraena zebra*, including mtDNA (left) and *RAG-1* and *RAG-2*. Rectangles represent inferred ancrestral haplotypes; circles are additional haplotypes, and small black-filled circles are inferred missing haplotypes. The size of the circle or square is proportional to the number of individuals sampled. Pie-chart shading denotes the oceanic region as demarked in Figure 4-1, with black for Indian Ocean, white for west/central Pacific, and grey for eastern Pacific Ocean.

Figure 4-3. Posterior probability estimates of bidirectional migration rates (for example, a value of 6 on the x-axis of the top figure indicates that on average 6 individuals move to

the Indian Ocean from the Central Pacific and 6 individuals move in the opposite

direction each generation), estimated for all three gene regions of *E. nebulosa*, among

biogeographic regions as delimited by the barriers of the Sunda Shelf and eastern Pacific

Barrier (see Figure 3-1). These estimates indicate equivalent and high levels of gene flow

among all ocean basins. The shape of the peak around the estimates indicates precision,

and the relative probability (Y-axis) indicates the relative support among migration
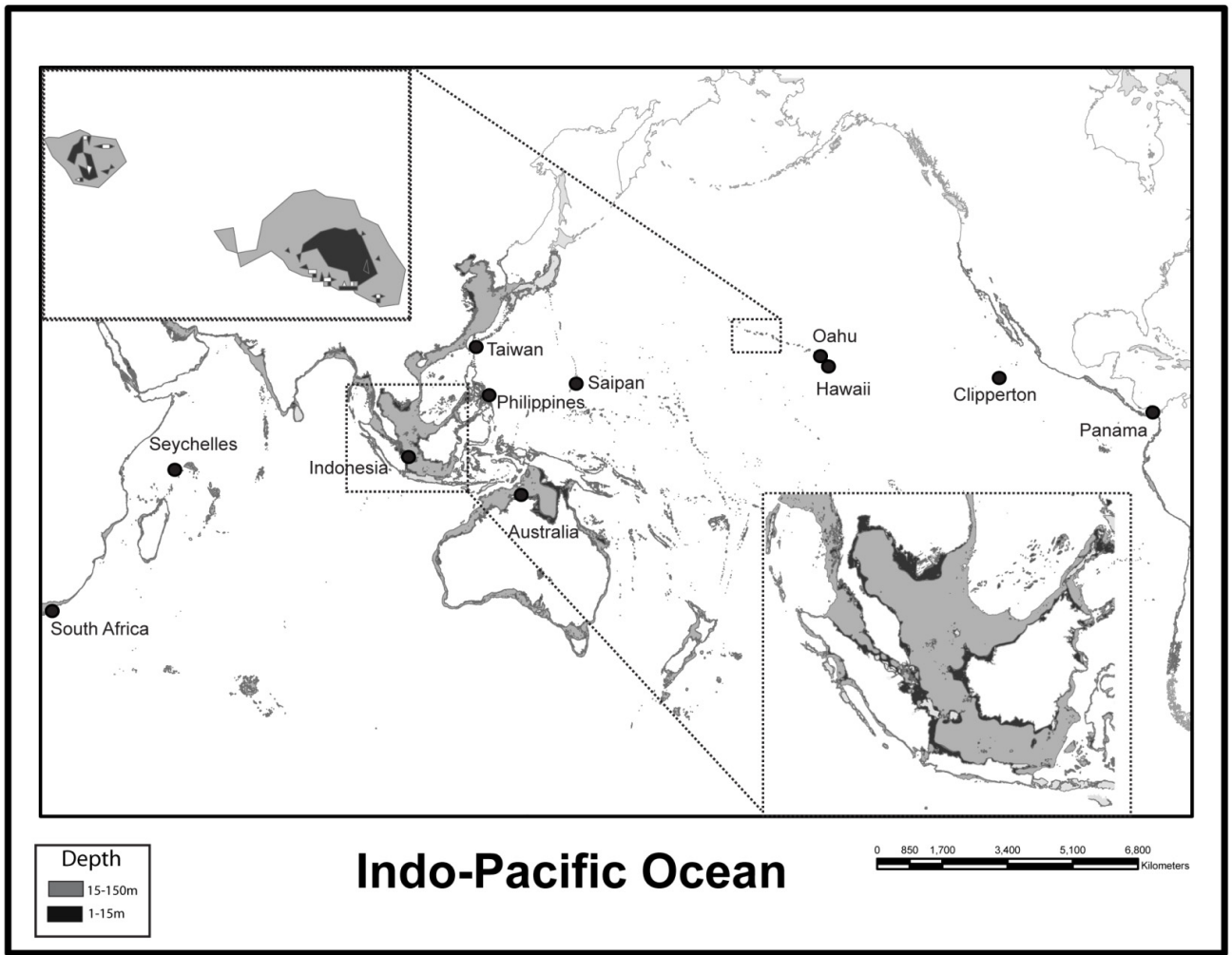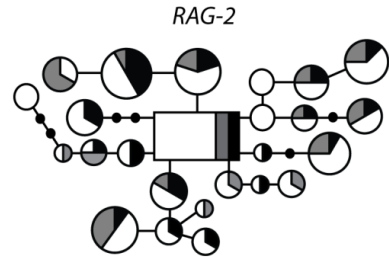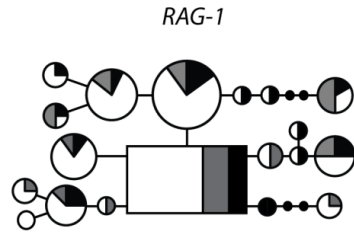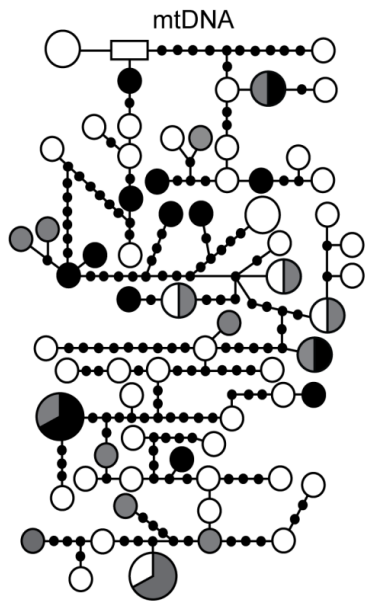
estimates sampled.

Indo-Pacific Ocean

Figure 4-1

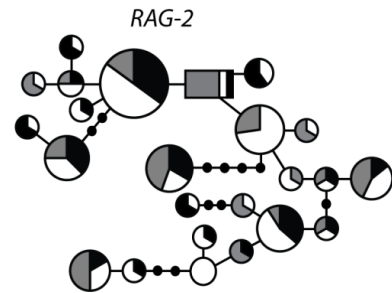*Echidna nebulosa*

mtDNA
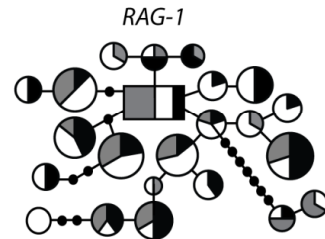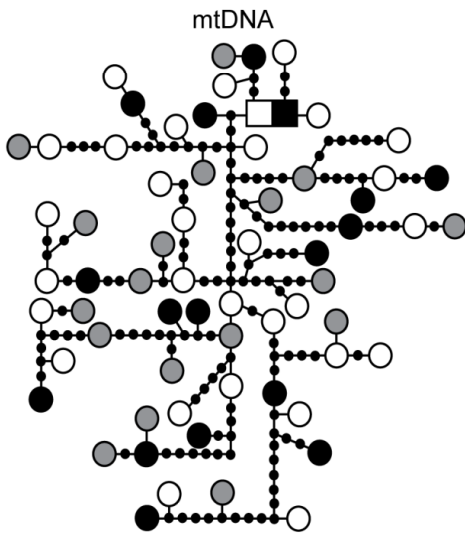
RAG-1

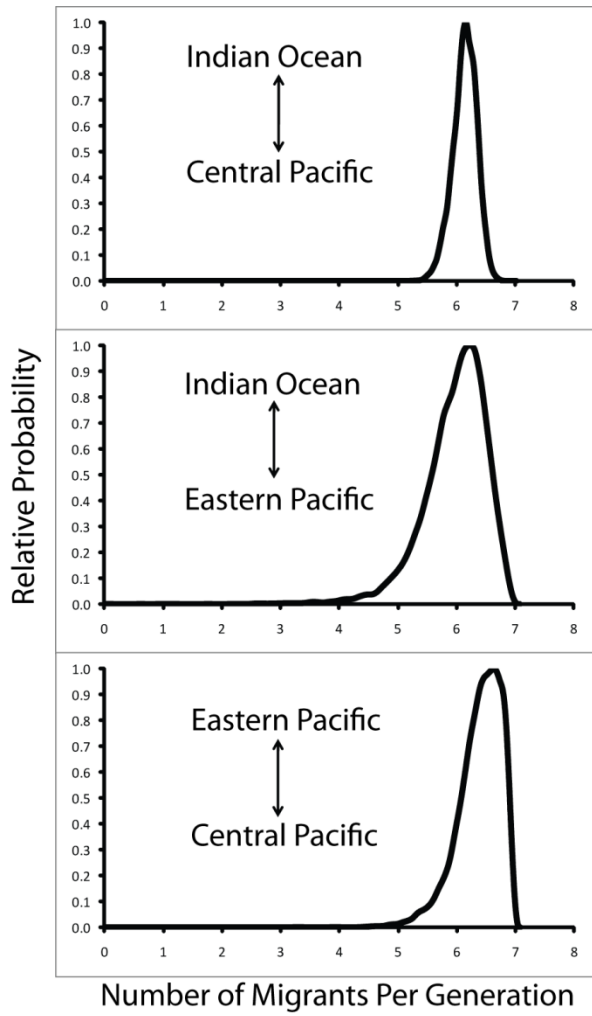RAG-2

Gymnomuraena zebra

mtDNA

RAG-1

RAG-2

Figure 4-2

Figure 4-3

Literature Cited

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Proceedings B 57, 289-300.

Bishop, R.E., Torres, J.J., Crabtree, R.E., 2000. Chemical composition and growth indices in leptocephalus larvae. Marine Biology 137, 205-214.

Böhlke, E.B., McCosker, J.E., Böhlke, J.E., 1989. Family Muraenidae. In: Böhlke, E.B. (Ed.), Fishes of the Western North Atlantic. Sears Foundation for Marine Research, New Haven.

Böhlke, E.B., Randall, J.E., 2000. A review of the moray eels (Anguilliformes: Muraenidae) of the Hawaiian Islands, with descriptions of two new species. Proceedings of the Academy of Natural Sciences of Philadelphia 150, 203-278.

Böhlke, E.B., Smith, D.G., 2002. Type Catalogue of the Indo-Pacific Muraenidae. Proceedings of the Academy of Natural Sciences of Philadelphia 152, 89-172.

Bonte, D., Baert, L., Lens, L., Maelfait, J.-P., 2004. Effects of aerial dispersal, habitat specialisation, and landscape structure on spider distribution across fragmented grey dunes. Ecography 27, 343-349.

Brothers, E.B., Thresher, R.E., 1985. Pelagic duration, dispersal, and the distribution of Indo-Pacific coral reef fishes. In: Reaka, M.L. (Ed.), Symposia Series for Undersea Research. NOAA Undersea Research Program, pp. 53-69.

Brändle, M., Brandl, R., 2001. Distribution, abundance and niche breadth of birds: scale matters. Global Ecology & Biogeography 10, 173-177.

Chatfield, B.S., Van Niel, K.P., Kendrick, G.A., Harvey, E.S., 2010. Combining environmental gradients to explain and predict the structure of demersal fish distributions. Journal of Biogeography 37, 593-605.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate genealogies. Molecular Ecology 9, 1657-1659.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7, 214.

Excoffier, L., Schneider, S., 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47-50.

Fasola, M., Canova, L., Foschi, F., Novelli, O., Bressan, M., 1997. Resource use by a Mediterranean rocky slope fish assemblage. Marine Ecology 18, 51-66.

Fu, Y.-X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915-925.

Gaughan, D.J., Potter, I.C., 1997. Analysis of diet and feeding strategies within an assemblage of estuarine larval fish and an objective assessment of dietary niche overlap. Fishery Bulletin 95, 722-731.

Harmelin-Vivien, M.L., 1979. Ichtyofaune des récifs coralliens de Tuléar (Madagascar): écologie et relations trophiques. Sciences. L'Université d'Aix-Marseille, Marseille.

Hawkins, B.A., Diniz-Filho, J.A.F., Jaramillo, C.A., Soeller, S.A., 2006. Post-Eocene climate change, niche conservatism, and the latitudinal diversity gradient of New World birds. Journal of Biogeography 33, 770-780.

Heino, J., Soininen, J., 2006. Regional occupancy in unicellular eukaryotes: a reflection of niche breadth, habitat availability or size-related dispersal capacity? Freshwater Biology 51, 672-685.

Hiatt, R.W., Strasburg, D.W., 1960. Ecological relationships of the fish fauna on coral reefs of the Marshall Islands. Ecological Monographs 30, 65-127.

Horne, J.B., van Herwerden, L., Choat, H.J., Robertson, D.R., 2008. High population connectivity across the Indo-Pacific: Congruent lack of phylogeographic structure in three reef fish congeners. Molecular Phylogenetics and Evolution 49, 629-638.

Hourigan, T.F., Reese, E.S., 1987. Mid-ocean isolation and the evolution of Hawaiian reef fishes. Trends Ecol. Evol. 2, 187-191.

Klanten, O.S., Choat, H.J., Van Herwerden, L., 2007. Extreme genetic diversity and temporal rather than spatial partitioning in a widely distributed coral reef fish. Marine Biology 150.

Kuhner, M.K., 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics Applications 22, 768-770.

Kuiter, R.H., 1998. Photo guide to fishes of the Maldives. Atoll Editions, Victoria, Australia.

Kuroki, M., Aoyama, J., Miller, M.J., Wouthuyzen, S., Arai, T., Tsukamoto, K., 2006. Contrasting patterns of growth and migration of tropical anguillid leptocephali in the western Pacific and Indonesian Seas.

Köckemann, B., Buschmann, H., Leuschner, C., 2009. The relationships between abundance, range size and niche breadth in Central European tree species. Journal of Biogeography 36, 854-864.

Leis, J.M., 1991. The pelagic stage of reef fishes: the larval biology of coral reef fishes. In: Sale, P.F. (Ed.), The ecology of fishes on coral reefs. Academic Press, New York, pp. 183-230.

Leis, J.M., Miller, J.M., 1976. Offshore distributional patterns of Hawaiian fish larvae. Marine Biology 36, 359-367.

Lester, S.E., Ruttenberg, B.I., 2005. The relationship between pelagic larval duration and range size in tropical reef fishes: a synthetic analysis. Proceedings of the Royal Society B 272, 585-591.

Lieske, E., Myers, R., 1996. Coral Reef Fishes: Caribbean, Indian Ocean and Pacific Ocean including the Red Sea. HarperCollins, London.

Mehta, R.S., 2008. Ecomorphology of the moray bite: relationship between dietary extremes and morphological diversity. Physiological and Biochemical Zoology 82, 90-103.

Mundy, B.C., 2005. Checklist of the fishes of the Hawaiian Archipelago. Bishop Museum Bulletin in Zoology 6, 1-704.

Myers, R.F., 1991. Micronesian reef fishes (2nd Edition). Coral Graphics, Guam.

Paillard, D., 2006. What drives the ice age cycle? Science 313, 455-456.

Pino, J., Vilà, M., Alvarez, N., Sequi, J.M., Guerrero, C., 2008. Niche breadth rather than reproductive traits explains the response of wetland monocotyledons to land-cover change. Applied Vegetation Science 12, 119-130.

Posada, D., 2008. jModelTest: Phylogenetic model averaging. Molecular Biology and Evolution 25, 1253-1256.

Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. Available from http://beast.bio.ed.ac.uk/Tracer.

Randall, J.E., 1967. Food habits of reef fishes of the West Indies. Studies of Tropical Oceanography 5, 665-847.

Randall, J.E., 1998. Zoogeography of shore fishes of the Indo-Pacific region. Zoological Studies 37, 227-268.

Randall, J.E., 2007. Reef and shore fishes of the Hawaiian Islands. Sea Grant College Program, University of Hawaii, Honolulu.

Reece, J.S., Bowen, B.W., Joshi, K., Goz, V., Larson, A., 2010. Phylogeography of two moray eels indicates high dispersal throughout the Indo-Pacific. Journal of Heredity 102, 1-12.

Reece, J.S., Bowen, B.W., Larson, A., in press. Molecular phylogenetics of moray eels (Murenidae) demonstrates multiple origins of a shell-crushing jaw (*Gymnomuraena*, *Echidna*) and multiple colonizations of the Atlantic Ocean. Molecular Phylogenetics and Evolution.

Riginos, C., Nachman, M.W., 2001. Population subdivision in marine environments: the contributions of biogeography, geographic distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. Molecular Ecology 10, 1439-1453.

Rosenberg, M.S., 2001. PASSAGE. Pattern Analysis, Spatial Statistics, and Geographic Exegesis. Version 1.1. Department of Biology, Arizona State University, Tempe, AZ.

Sampayo, E.M., Franceschinis, L., Hoegh-Guldberg, O., Dove, S., 2007. Niche partitioning of closely related symbiotic dinoflagellates. Molecular Ecology 16, 3721-3733.

Scharf, F.S., Juanes, F., Rountree, R.A., 2000. Predator size-prey size relationships of marine fish predators: interspecific variation and the effects of ontogeny and body size on trophic-niche breadth. Marine Ecology Progress Series 208, 229-248.

Sommer, C.W., Schneider, W., Poutiers, J.M., 1996. FAO species identification field guide for fishery purposes. The Living Marine Resources of Somalia, Rome.

Soons, M.B., Messelink, J.H., Jongejans, E., Heil, G.W., 2005. Habitat fragmentation reduces grassland connectivity for both short-distance and long-distance wind-dispersed forbs. Journal of Ecology 93, 1214-1225.

Swihart, R.K., Gehring, T.M., Kolozsvary, M.B., 2003. Responses of 'resistant' vertebrates to habitat loss and fragmentation: the importance of niche breadth and range boundaries. Diversity and Distributions 9, 1-18.

Swihart, R.K., Lusk, J.J., Duchamp, J.E., Rizkalla, C.E., Moore, J.E., 2006. The roles of landscape context, niche breadth, and range boundaries in predicting species responses to habitat alteration. Diversity and Distributions 12, 277-287.

Templeton, A.R., 2006. Population genetics and microevolutionary theory. John Wiley & Sons, Hoboken, N.J.

Thompson, K., Gaston, K., Band, S., 1999. Range size, dispersal and niche breadth in the herbaceous flora of central England. Journal of Ecology 87, 150-155.

Thresher, R.E., 1984. Reproduction in Reef Fishes. T.F.H. Publications, Neptune City, N.J.

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

Young, A., Boyle, T., Brown, T., 1996. The population genetic consequences of habitat fragmentation for plants. Trends Ecol. Evol. 11, 413-418.

Yukihira, H., Shibuno, T., Hashimoto, H., Gushima, K., 1994. Feeding habits of moray eels (Pisces: Muraenidae) at Kuchierabu-jima. Journal of the Faculty of Applied Biological Science 33, 159-166.

CHAPTER 5

PAGES 104-11

CONCLUSIONS AND FUTURE WORK

The goals of this work are to describe the biodiversity of moray eels on two levels: the phylogenetic structure, evolutionary history and relationships among species, and secondly, the geographic structuring of populations within species. My results confirm higher-level taxonomy of subfamilies, but generic relationships often do not reflect evolutionary history. The genus *Echidna*, previously united by a shared durophagous cranial morphology, actually represents parallel evolution of durophagy in at least three lineages, making the genus polyphyletic. I also show that the evolutionary histories of moray eel species in the Atlantic and Indo-Pacific are intertwined and represent several exchanges throughout the last 50 million years. At the interspecific level, I tested the prediction that long pelagic larval durations (PLD) in morays would confer high levels of gene flow and low population structure. This pattern was elucidated in four species, with implications that the highly dispersive larva eradicates any influence of adult niche breadth on phylogeographic structure among populations.

The taxonomy and classification of understudied groups is often based solely on external morphological and basic skeletal features. This is problematic when morphological features that are shared among taxa do not reflect shared ancestry. My initial phylogenetic work appeared to support rampant polyphyly of named taxa based on museum specimens, but was later shown to document high levels of misidentification of specimens in museum collections. My work on moray eels substantiated the hypotheses of Nelson (Nelson, 1966) separating the snakemorays (Uropterygiinae) from the Muraeninae. This result confirms similar surveys by Nelson (1966) and Loh et al. (2008), and is subject to a more exhaustive sampling of species in both subfamilies.

The primary taxonomic revisions resulting from this work relate to the genus *Echidna*. This genus is united by a durophagous condition, and the four species sampled here represent three independent origins of durophagy (Figure 2-2). The genus is formally recognized by a type specimen of *Echidna nebulosa*, a species that I sampled. Therefore, *Echidna nebulosa* and any other species more closely related to *E. nebulosa* than to any species of *Gymnothorax* would retain the genus *Echidna*. Until more detailed morphological analyses can be made, I would suggest that other members of the genus sampled here (*E. rhodochilus*, *E. leucotaenia*, and *E. polyzona*) revert to the genus *Gymnothorax*. It is not surprising that the genus *Gymnothorax* is paraphyletic, as taxonomists have used this genus as a catch-all, and the genus currently forms approximately 60% of all of Muraenidae. The simplest taxonomic revision that would reflect patterns of ancestry would be to subsume the genera *Enchelynassa*, *Rhinomuraena*, and *Enchelycore* within the genus *Gymnothorax*. With these revisions, genera *Echidna* and *Gymnothorax* would be sister taxa based on the sampling contained in this dissertation.

Many lineages of reef fishes show separate radiations in the Atlantic and Indo-Pacific Oceans (Barber and Bellwood, 2005; Bernardi et al., 2004; Bowen et al., 2006b; Muss et al., 2001). My sampling of Atlantic lineages was limited to five species, but even at this level clearly demonstrated multiple exchanges between the Atlantic and Indo-Pacific over varied timescales, approximately 21, 16, and 2.9 million years ago. The overall structure of my phylogenetic reconstruction and time calibrations confirms the age of Muraenidae at approximately 56 My, with most lineage accumulation occurring during the Miocene.

Finally, this work evaluates effects of pelagic larval duration on phylogeographic structure. Because most reef fish accomplish dispersal only as larvae floating or swimming in open ocean currents (as opposed to the reef habitat of the adult form), the length of time that larvae can exist in the pelagic realm was thought to be predictive of the spatial scale over which individuals could disperse (Jones et al., 2005; Taylor and Hellberg, 2005). Until the late 2000s, this rule was relatively unchallenged in the literature. Recent surveys have challenged this dogma and emphasized the importance of complicating factors (e.g., Bowen et al., 2006a). Weersing and Toonen (2009) analyzed a dataset of fish and invertebrates and noted only a weak correlation between pairwise F_{ST} measures and average pelagic larval duration in some groups. I suggest in Chapter 4 that a pelagic larval duration in excess of 80 days is sufficient to erase patterns of isolation by distance over evolutionary timescales. A linear decrease of pairwise F_{ST} is therefore not expected for pelagic larval durations exceeding ~80 days. The Weersing and Toonen (2009) survey included five species of fish with average larval durations of more than 80 days. The specific results from these taxa are not reported, but overall they observed a weak but significantly negative correlation with PLD and population genetic structure. My results do not address the ability of PLD to predict population structure across a continuum of short to long larval durations, because all of the morays sampled in my study have uniformly long larval durations. I confirm the hypothesis that species with long larval durations will have very low population genetic structure, and this finding is a prediction of the Weersing and Toonen study (see Figure 1 in Weersing and Toonen, 2009). In the case of widely-distributed and highly dispersive taxa such as moray eels, extended pelagic larval durations do confer high levels of connectivity among

populations, even across 2/3rds of the planet. My work on *Gymnothorax undulatus*, *Gymnothorax flavimarginatus*, *Gymnomuraena zebra*, and *Echidna nebulosa*, in combination with phylogenetic analyses of 40 other species, reveal this pattern across lineages separated by greater than 20 million years of independent evolutionary histories. I also confirm the prediction that if larval dispersal capabilities do prevent isolation by distance, then similar low levels of population structure should be observed across taxa despite enormous differences in adult niche breadth and adult habitat fragmentation. The patterns of widespread connectivity in *G. zebra* and *E. nebulosa* demonstrate that larval dispersal capability can supersede factors that would otherwise cause greater geographic structuring of gene lineages.

Future Work

This work documents the evolutionary history of 44 out of 202 species of moray eels (~22%). These species were chosen in an attempt to encompass most of the evolutionary history of Muraenidae, but much work remains to be done. The interpretations that I draw from Chapter 2 are robust to additional sampling, with the exception of the monophyly of the subfamilies Muraeninae and Uropterygiinae. Additional sampling of the morphologically unique genera *Strophidon* and *Channomuraena*, and more exhaustive sampling of the large, catch-all genus *Uropterygius* (21 of the 36 species in Uropterygiinae) would help to resolve the monophyly of the Uropterygiinae. It is unclear which members of the genus *Uropterygius* would be most useful to sample, but I would suggest at least eight additional species (for a total of ten) to capture most of the evolutionary history of the group. The 41 species of Muraeninae sampled include members of every major morphotype with the exception of

two species of the genus *Strophidon*, however the deepest divergences within Muraeninae are thought to be between the genus *Gymnothorax* and the monotypic genus *Gymnomuraena*, which I did sample. Based on the available evidence, the phylogenetic sampling in Chapter 2 likely encompassed the deepest divergences in Muraenidae, but potentially missed the deepest split within Uropterygiinae. It is thus possible that future sampling could render Uropterygiinae paraphyletic with respect to Muraeninae.

A more complete sampling of species would also allow for precise estimates of lineage accumulation rates and could better describe recent patterns of diversification throughout the evolutionary history of moray eels. All of the major morphotypes within the Muraenidae were included in the phylogenetic sampling of Chapter 2, with the exception of two species of *Channomuraena* and one species of *Strophidon*. Additional sampling might identify more recent cladogensis, but my current sampling of 44 of approximately 200 species is likely to capture the deepest divergences within the group, in particular given the fact that this sampling included members of the two subfamilies that define the largest differences, and putatively the deepest divergences, in all of Muraenidae.

I consider my phylogeographic sampling within the four species (*Gymnothorax undulatus*, *Gymnothorax flavimarginatus*, *Gymnomuraena zebra*, and *Echidna nebulosa*) definitive with respect to demonstrating ocean-wide gene flow and genetic homogeneity. Additional sampling, especially of the south Pacific, would at most identify regional fragmentation of areas not sampled, but such findings would not reverse the conclusion that the widely separated populations actually sampled have high average levels of gene flow among them. All morays with extended larval durations are capable of maintaining

geographic-genetic homogeneity over millions of years. There is evidence that on shorter timescales, corresponding with 100,000 cycles of sea level rise and fall (Paillard, 2006), local population genetic structure can evolve and potentially produce cladogenesis. Over longer evolutionary timescales, those lineages either homogenize or each acquires ocean-wide distributions that are typically indistinguishable from panmixia. In other words, evolutionary timescales show evidence of widespread gene flow, but on shorter ecological timescales (tens to a few hundred thousand years), geographic population structure can and does evolve. It will be interesting to see if comparable phylogeographic patterns occur in species of Uropterygiinae, which I did not sample for phylogeographic structure. I would suggest that future studies begin with phylogeographic study of *Uropterygius macrocephalus*, a small and commonly observed species that can be captured by hand in shallow water and is distributed throughout the Indo-Pacific from the western coast of the Americas to the east and as far west as the Seychelles. These species are restricted to shallow water between zero and 14m (Myers, 1999), but restricted distributions by depth do not appear to affect connectivity of populations (Chapter 4).

Literature Cited

Barber, P.H., Bellwood, D.R., 2005. Biodiversity hotspots: evolutionary origins of biodiversity in the wrasses (*Halichoeres*: Labridae) in the Indo-Pacific and new world tropics. Molecular Phylogenetics and Evolution 35, 235-253.

Bernardi, G., Bucciarelli, G., Costagliola, D., Robertson, D.R., Heiser, J.B., 2004. Evolution of coral reef fish *Thalassoma* spp. (Labridae). 1. Molecular phylogeny and biogeography. Marine Biology 144, 369-375.

Bowen, B.W., Bass, A.L., Muss, A., Carlin, J., Robertson, D.R., 2006a. Phylogeography of two Atlantic squirrelfishes (Family Holocentridae): exploring links between pelagic larval duration and population connectivity. Marine Biology 149, 899-913.

Bowen, B.W., Muss, A., Rocha, L.A., Grant, W.S., 2006b. Shallow mtDNA coalescence in Atlantic Pygmy Angelfishes (Genus *Centropyge*) indicates a recent invasion from the Indian Ocean. Journal of Heredity 97, 1-12.

Jones, G.P., Planes, S., Thorrold, S.R., 2005. Coral reef fish larvae settle close to home. Current Biology 15, 1314-1318.

Loh, K.-H., Chen, I.-S., Randall, J.E., Chen, H.-M., 2008. A review and molecular phylogeny of the moray eel subfamily Uropterygiinae (Anguilliformes: Muraenidae) from Taiwan, with description of a new species. Raffles Bulletin of Zoology 19, 135-150.

Muss, A., Robertson, D.R., Stepien, C.A., Wirtz, P., Bowen, B.W., 2001. Phylogeography of *Ophioblennius*: the role of ocean currents and geography in reef fish evolution. Evolution 55, 561-572.

Myers, N., 1999. Micronesian Reef Fishes: A Comprehensive Guide to the coral reef
fishes of Micronesia, 3rd revised and expanded edition. Coral Graphics,
Barrigada, Guam.

Nelson, G.J., 1966. Gill arches of teleostean fishes of the order Anguilliformes. Pacific
Science 20, 391-408.

Paillard, D., 2006. What drives the ice age cycle? . Science 313, 455-456.

Taylor, M.S., Hellberg, M.E., 2005. Marine radiations at small geographic scales:
speciation in neotropical reef gobies (*Elacatinus*). Evolution 59, 374-385.

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