

**Evolutionary patterns of alternative life histories in fishes:
assessing marine-freshwater and aquatic-terrestrial movement**

**A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

**Joel Benjamin Corush
December 2018**

Copyright © 2018 by Joel Benjamin Corush
All rights reserved.

DEDICATION

To my friends, family and colleagues.

ACKNOWLEDGEMENTS

I want to thank my advisor, Benjamin Fitzpatrick, who allowed me the freedom to peruse the questions that I was interested in and helped me to develop my own research program. I would also like to thank my committee, Graciela Cabana, Ben Keck and Dan Simberloff. You all have made me become a better scientist by helping me think critically about the questions I ask, the methods I use, and how I interpret data.

I would not have been able to conduct fieldwork without the generous help of many people who spent their time, effort and money to help me with logistics of field work, but also provided a warm and welcoming environment. In Guam, Brent Tibbatts from Guam Department of Agriculture, Division of Aquatic and Wildlife Resources, as well as Drs. Daniel Lindstrom, Kate Moots and Curt Fiedler from University of Guam. In Taiwan, Dr. Jen-Chieh Shiao and the members of his lab: Dr. Ni-Na Chang, Mr. Jyun Liao, Mr. Pin-Ren Fang, 黃文謙, 袁潋晴, 鄭力綺 and 呂翰駁 from National Taiwan University. In China, Dr. Jie Zhang and Mr. Feng Lin, from the Key Laboratory of Animal Ecology and Conservation Biology Institute of Zoology, Chinese Academy of Sciences. In Japan Dr. Zdenek Lajbner from the Okinawa Institute of Science and Technology. Also Yukitoshi Katayama (片山侑駿) and Dr. Atsushi Ishimatsu students, Mai Van Hieu and Tran Xuan Loi who provided samples from locations that I was unable to go to.

Molecular work would not have gone so smoothly without the tremendous assistance from Travis C Glenn, Natalia J Bayona-Vasquez and Troy J Kieran at University of Georgia.

The faculty, staff and students of the Ecology and Evolutionary Biology Department. Evin Carter, Todd Pierson, Zach Marion, Orlando Schwery, Jeremy Beaulieu, Brian O'Meara, James Fordyce, Monica Papes, Darrin Hulsey, Mike Harvey and Elizabeth Wolfe for all their help with different aspect of my research. And Jeremiah Henning, Courtney Gorman, Angela Chuang, Sharon Munoz, Sam Borstein, Cedric Landerer, Cassie Dresser-Briggs, Christopher Peterson, Diane Le Bouille, Lisa Cantwell, Melquisedec Gamba-Rios, Tyson Paulson, Jonathan Dickey, Shelby Scott and Hailee Korotkin for all that I learned from our discussions, seminars as well as their general support during my Ph.D. Thank you to the EEB office staff who helped organize my funding, travel and general logistics of graduate school: Marva Anderson, Lisa Boling, Karin Grindall and Janice Harper.

And lastly, to my family who has supported and encouraged me over the years. Regina and Chuck Corush, Jenna Corush, and Steve, Mara, Sophia, Liam and Juliet Corush, thank you.

I would also like to acknowledge my funding sources: UTK-Ecology and Evolutionary biology department research funds, UTK Chancellors fund, UTK McClure scholarship,

Sigma-Xi Grants-in-Aid of Research, and The National Science Foundation (awards #1414902 and #1614100).

ABSTRACT

Transitions between habitats driven by variation in physiological requirements at different stages of ontogeny have a wide range of effects on individuals, populations, and species as a whole. Where and when individuals can reproduce, how far they can disperse, and their physical ability to withstand particular environments affect gene flow and population size, which in turn may affect the potential for speciation and extinction. Explaining how individual-level behavior can alter diversification of species is a vast and challenging endeavor evolutionary biologists face. For my dissertation I address diadromous and amphibious behaviors in fishes. Diadromous fishes migrate between marine and freshwater systems and amphibious fishes move between aquatic and terrestrial habitats. Both behaviors restrict individuals' movements and are thought to be important in population and species level patterns. First, I use phylogenetic comparative methods in order to test the hypothesis that diadromy is an intermediate state between completely marine and freshwater life styles. I find that while in some cases diadromy seems to act as an evolutionary intermediate, it is also a state where lineages diversify before returning back to their ancestral environment. This suggests that diadromy is an important behavior in the diversification of some groups of fishes. Second, I use comparative population genetics to see if the high diversification rates found in diadromous fishes is associated with variation in population differentiation and genetic diversity. I find that there is no particular association with migratory life history and population structure in fishes. Migration life history does not predict levels of differentiation between populations nor genetic diversity within populations. Finally, I use population genetic tools to see if the terrestrial restrictions of amphibious behavior affect population structure. Assessing a mudskipper from East Asia, *Periophthalmus modestus*, I find that movement and connectivity of populations is driven by connectivity of land. Large, diverse continental island populations are likely the result of vicariance and small, less diverse oceanic island populations are likely the result of rare colonization events. These findings show that habitat restrictions are important in shaping patterns of biodiversity in fishes.

TABLE OF CONTENTS

INTRODUCTION	1
References.....	3
CHAPTER I THE EVOLUTIONARY SIGNIFICANCE OF DIADROMY AS A TRANSITIONAL STATE BETWEEN MARINE AND FRESHWATER IN FISHES... .	5
Abstract.....	6
Introduction.....	7
Methods	10
Results	12
Discussion.....	13
Conclussion.....	18
References.....	18
Appendix A.....	24
Appendix B.....	30
Appendix C	34
CHAPTER II GENETIC DIVERSITY AND POPULATION STRUCTURE OF MARINE, CATADROMOUS, ANADROMOUS, AMPHIDROMOUS AND FRESHWATER FISHES.....	35
Abstract.....	36
Introduction.....	37
Methods	41
Results	43
Discussion.....	44
References.....	48
Appendix A.....	60
Appendix B.....	67
CHAPTER III THE EFFECTS OF AN AMPHIBIOUS LIFE HISTORY ON THE POPULATION STRUCTURE OF A MUDSKIPPER (PERIOPHTHALMUS MODESTUS) IN EAST ASIA.....	81
Abstract.....	82
Introduction.....	83
Methods	85
Results	89
Discussion.....	90
References.....	93
Appendix A.....	98
Appendix B.....	105
CONCLUSION.....	118
VITA	119

LIST OF TABLES

Table 1.1 Output parameters for key model in MuSSE and fitDiscrete analyses	29
Table A1.1 Model constraints for all MuSSE analyses	30
Table A1.2 Output parameters for optimal model in MuSSE analyses with “unknown” species in either freshwater or marine	31
Table A1.3 Models constraints for HiSSE and BiSSE analysis	32
Table 2.1 Heterozygosity and max length by species	62
Table 2.2 Heterozygosity models	64
Table 2.3 Isolation by distance tables	65
Table 2.4 Life history models	66
Table A2.1 Populations locations	69
Table A2.2 Heterozygosity phylogenetic linear regressions	79
Table A2.3 Euclidian and shortest path distances	80
Table 3.1 Overall and populations specific Heterozygosity (H_e)	101
Table 3.2 Pairwise F_{st}	101
Table 3.3 Mantel and partial mantel test of genetic and geographic distances	102
Table 3.4 ABC model selection	103
Table 3.5 ABC parameter estimates	104

LIST OF FIGURES

Figure 1.1 Relative distribution of diadromy in fishes.....	24
Figure 1.2 Illustrated hypotheses of the evolution of diadromy.	26
Figure 1.3 Posterior probability distributions for MuSSE output paramaters.....	27
Figure 1.4 Simulated distribution of marine, freshwater, and diadromous fishes..	28
Figure 2.1 IBS slopes arranged by life history.	60
Figure 2.2 Distribution of Heterozygosity.....	60
Figure 2.3 Distribution of maximum length.....	61
Figure A2.1 Reconstructed phylogeny used for IBD analyses.....	67
Figure A2.2 Reconstructed phylogeny used for gene diversity analyses.....	68
Figure 3.1 Population locations of <i>Periophthalmus modestus</i> across the South and East China seas	98
Figure 3.2 STRUCTURE population assignment plots.	99
Figure 3.3 Migration and dissimilarity heatmaps.	100
Figure 3.4 Genetic variation plotted against the three forms of geographic distance.....	100
Figure A3.1 Histogram of frequency of number of loci per individual sample e.....	105

LIST OF ATTACHMENTS

Attachment 1.1 Species character state assignment (SuppTable1.xls).	34
---	----

INTRODUCTION

Major transitions between habitats are seen across the tree of life. Transition between aquatic and terrestrial (Shubin *et al.* 2006; Uhen 2007), flightlessness and flight (Garner *et al.* 1999; Dudley *et al.* 2007), and marine and fresh water (Logares *et al.* 2008; Bloom *et al.* 2013) all require major behavioral, physiological and morphological changes, yet each has occurred numerous times across the tree of life. Alternative life histories that require an organism to spend different portions of its life in fundamentally different habitats are found in many taxonomic groups including amphibians, fishes, insects, and crustaceans. Biphase organisms are often used as models for understanding major evolutionary transitions (e.g., Linzey 2012), but the persistence of biphase groups suggests that biphase life histories can be evolutionarily stable, not just transient phases of an inevitable transition. Explaining the persistence of these transitional states is an important challenge for evolutionary biologists as these transitions may allow species to move into previously unoccupied niches. Among fishes, two of the most dramatic changes in habitat types are the transitions between marine and freshwater and between aquatic and terrestrial habitats. Numerous species with alternative life histories parallel each of these evolutionary transitions.

In fishes, transitions between marine and freshwater life histories have occurred many times at the individual, population and species level. While some organisms tolerate a wide range of salinity (Griffith 1974), most prefer consistency for their entire lives. Diadromous organisms, or those with scheduled transition between freshwater and marine environment at specific life stages (Myers 1949, McDowall 1988), require substantial shifts in physiology during transitions. Movement between hypo- and hyperosmotic conditions requires change in osmoregulation and buoyancy regulation. The migration itself also requires a large requirement of energy and increases the risk of death. Together these factors equate to a large potential cost associated with this behavior.

Amphibious fishes are also of interest as they require the use of both aquatic and terrestrial habitats at some points during an individual's life cycle. Many fish can persist on land for a period of time (Sayer and Davenport 1991), but few species are obligately amphibious. A fish's locomotion and breathing require numerous morphological changes when it steps out of water (Kawano and Blob 2013). The potential for dehydration, increased risk of predation due to terrestrial exposure, and the desiccation of eggs, all pose serious threats to amphibious fishes.

My dissertation addressed questions about the effects of these two different biphase lifestyles on the diversity and distribution of fishes. In my first chapter, I examine whether there is an evolutionary directionality in transitions in and out of these different lifestyles. I address previous hypotheses that suggest diadromy is an evolutionary stepping stone between completely marine and completely freshwater life histories. I used comparative phylogenetic methods to show that, while diadromy may act as an intermediate between marine and freshwater, it is also a source of species diversity leading to an increase in species diversification rates. In some cases it follows the

evolutionary trajectory proposed by Gross *et al.* (1988) and Dodson *et al.* (2009), but in many cases, it does not and lineages revert back to their ancestral environments.

In my second chapter, we test to see if the diversification rates observed in the first chapter correspond with population differentiation with the expectation that behaviors that result in high speciation rate do so as a function of strong population subdivision. We looked at levels of population differentiation across fishes with five different life histories but found no significant variation between marine, freshwater, anadromous, catadromous and amphidromous fishes.

In my third chapter I assess the second alternative life history in an amphibious mudskipper distributed in the South and East China Seas, *Periophthalmus modestus*. I use a variety of population genetic analyses coupled with spatial and coalescent analyses to suggest that movement is restricted to mudflats and costal habitat, resulting in low levels of across water dispersal. Population connectivity in mudskippers depends on continuous coastal areas.

This work addresses macroevolutionary and microevolutionary questions by utilizing a variety of phylogenetic and population genetic methods to provide a holistic view of the evolutionary consequences of movement between environments as part of alternative life histories. I show that diversification and population structures are both affected by these life histories.

References

- Bloom, D. D. and N. R. Lovejoy. The evolutionary origins of diadromy inferred from a time-calibrated phylogeny for Clupeiformes (herring and allies). *P Roy Soc B-Biol Sci.* 2014;281(1778):2013-2081.
- Dodson, J. J., J. Laroche, and F. Lecomte. Contrasting evolutionary pathways of anadromy in euteleostean fishes. In *Challenges for diadromous fishes in a dynamic global environment*. *Am Fish S S.* 2009;69:63-77.
- Dudley, R. et al. Gliding and the functional origins of flight: biomechanical novelty or necessity? *Annu Rev Ecol Evol S.* 2007;38:179-201.
- Garner, J.P., Taylor, G.K. and Thomas, A.L., 1999. On the origins of birds: the sequence of character acquisition in the evolution of avian flight. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1425), pp.1259-1266.
- Griffith, R.W., 1974. Environment and salinity tolerance in the genus *Fundulus*. *Copeia*, pp.319-331.
- Gross, M. R., R. M. Coleman, and R. M. McDowall. Aquatic productivity and the evolution of diadromous fish migration. *Science* 1988;239(4845):1291-1293.
- Kawano, S.M. and Blob, R.W., 2013. Propulsive forces of mudskipper fins and salamander limbs during terrestrial locomotion: implications for the invasion of land. *Integrative and comparative biology*, 53(2), pp.283-294.
- Linzey, D.W., 2012. *Vertebrate biology*. JHU Press.
- Logares, R., Bråte, J., Bertilsson, S., Clasen, J.L., Shalchian-Tabrizi, K. and Rengefors, K., 2009. Infrequent marine–freshwater transitions in the microbial world. *Trends in microbiology*, 17(9), pp.414-422.
- McDowall, R. M. *Diadromy in Fishes: Migrations Between Freshwater and Marine Environments*. Timber Press, Portland, OR. 1988.
- Myers, G. S. Usage of anadromous, catadromous and allied terms for migratory fishes. *Copeia*. 1949;2:89-97.
- Sayer, M.D.J. and Davenport, J., 1991. Amphibious fish: why do they leave water?. *Reviews in Fish Biology and Fisheries*, 1(2), pp.159-181.
- Shubin, N. H., E. B. Daeschler, and F. A. Jenkins. The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature*. 2006;440(7085):764-771.

Uhen, M.D., 2007. Evolution of marine mammals: back to the sea after 300 million years. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology: Advances in Integrative Anatomy and Evolutionary Biology*, 290(6), pp.514-522.

CHAPTER I
THE EVOLUTIONARY SIGNIFICANCE OF DIADROMY AS A
TRANSITIONAL STATE BETWEEN MARINE AND FRESHWATER
IN FISHES

Abstract

Background: Many organisms exhibit life history stages that require movement between different environments during the course of their lives (i.e. aquatic v. terrestrial, flightlessness v. flight, marine v. freshwater). These developmental shifts parallel dramatic evolutionary shifts that require physiological, anatomical and behavioral changes for survival and reproduction. Diadromy (scheduled movement between marine and freshwater) has been characterized as an evolutionary transitional state between different environments, implying that diadromous lineages are evolutionarily transient. Some hypotheses even propose that diadromy is a necessary transitional state, implying that direct evolutionary shifts from marine to freshwater (or vice versa) are extremely unlikely. Here I evaluate these predictions using comparative methods and the most up-to-date phylogeny of fishes.

Results: Models including direct evolutionary transitions between freshwater and marine were supported over those with diadromy as a necessary intermediate. However, estimated evolutionary transition rates from diadromous lifestyles to marine or freshwater are 10-100 times higher than transition rates directly between marine and freshwater. Additionally, transitions from marine or freshwater into diadromous lifestyles are rare, but the rate of diversification in diadromous fishes is ten times higher than in marine or freshwater taxa.

Conclusion: Diadromous life histories evolve relatively rarely in fishes. However, diadromous lineages have high rates of diversification compared to marine or freshwater species, and give rise to marine and freshwater specialists in addition to diadromous descendants. Although diadromy is not a necessary evolutionary intermediate, diadromous lineages are sources of both marine and freshwater diversity. These results support the interpretation of diadromy as an important, but not necessary, intermediate state that ontogenetically recapitulates the transition between marine and freshwater. Diadromous lineages tend to be more transient than marine or freshwater lineages. This evolutionary instability of diadromous lineages is counteracted by their relatively high diversification rates. These findings highlight the importance of integrating the dynamics of diversification and major evolutionary transitions for understanding macroevolutionary patterns.

Introduction

Major transitions between different habitat types are seen across the tree of life. Transitions between water and land [1], flightlessness and flight [2], and marine and freshwater [3] all require major behavioral, physiological and morphological changes, yet each transition has occurred numerous times [1-7]. Explaining the high frequency of these transitions is an important challenge for evolutionary biologists, as these transitions allow species to move into niches not previously occupied by their lineage. Life histories in which these major transitions occur ontogenetically are found in many taxonomic groups, including salamanders that reproduce in water and live on land; fishes, crabs and snails that migrate between marine and freshwater; sea turtles that reproduce on land and spend the rest of their lives in the ocean; and many semi-aquatic insects, such as dragonflies, that have an aquatic larval stage but a terrestrial adulthood.

Among fishes, one of the most dramatic changes in habitat type is between marine and freshwater. Diadromy, or the migration between marine and freshwater during a particular life stage [4,6], has been hypothesized as an evolutionary transition between completely marine and completely freshwater lifestyles [5,8,9]. These transitions require numerous physiological changes in osmotic and ionic regulation [10]. While many euryhaline fishes can move across a wide gradient of salinity with regularity, diadromous species move between fresh and marine environments only at specific life stages [4,6,11]. This scheduled movement between environments is of ecological and evolutionary interest, as it is seen in only about 0.8% of the 32,000 species of ray-finned fishes, yet it is found in around 8% of the nearly 500 families of fishes [6] (Fig. 1.1). In addition, diadromous behavior is observed in other parts of the tree of life including in many invertebrates such as snails [7], prawns [12], and crabs [13,14].

Multiple hypotheses have been proposed as to the evolutionary significance of diadromy [5,8,9,15]. Many studies focus on the transitions in and out of diadromy and propose that diadromy acts as an intermediate stage between completely marine and completely freshwater life histories. While the proposed mechanisms range from resource availability for adults based on latitudinal variation in net primary productivity of environments [8] to moving into a safe site for spawning [9], each of these hypotheses suggests an evolutionary trajectory along these lines: a marine species has individuals or lineages that venture into freshwater (be it during adulthood or spawning) and over time, lineages stop returning to marine systems, becoming completely freshwater due to some advantage(s) of the new environment. Similar arguments have been constructed in the opposite direction for transitions from freshwater to marine [5,8]. The natural world provides many examples of evolutionary transitions from diadromous to solely freshwater lifestyles in sticklebacks (*Gasterosteidae*), lampreys (*Petromyzontiformes*), freshwater eels (*Anguillidae*), salmonids (*Salmonidae*) and alewives (*Clupeidae*) [6,16,17]. There are also examples of diadromous species where individuals spend their entire lives in marine systems and phylogenies that support an evolutionary trend towards marine lifestyles within the lineage [6,16,18,19]. While some of these examples support

the directionality in the previously proposed hypotheses about the evolution of diadromy [3,17], examples of evolution from marine to diadromy and back to marine do not [3,18,19]. There are also some putative examples of transition directly between marine and freshwater [20-23].

In addition to the transitions between life histories, speciation and extinction rates can be incorporated into the hypotheses of diadromy as an evolutionary intermediate. If diadromy is solely a transitional form, it would be expected that there would be little speciation or extinction of diadromous lineages. This simple prediction is contradicted by examples of clades with many diadromous sister species (such as salmonids, eels, and gobies). However, the question remains whether diadromous lineages tend to have different speciation or extinction rates than marine or freshwater lineages. Drastic differences in either rate would suggest that the life history of diadromy is changing the net diversification rate (net diversification rate = speciation rate – extinction rate) of fishes once the behavior has evolved. Diadromous behavior has been proposed to have two seemingly conflicting effects on genetic structure and components of net diversification. On one hand, diadromous migration between different rivers is expected to decrease population structure compared to completely freshwater species which cannot move between river systems via the ocean. This should decrease the risk of extinction due to unfavorable habitat changes or stochastic events that would wipe out a completely freshwater species that is fully restricted to that freshwater system. At the same time diadromy could decrease the rate of speciation due to increased gene flow between isolated river systems via migration into and back out of the ocean. On the other hand, diadromous migrations are expected to increase the geographic range of a species compared to completely freshwater taxa, potentially increasing isolation-by-distance and changing the selection pressure put on a population as it expands into a new habitats. This might increase rates of ecological speciation via local adaptation or vicariant speciation via landlocking or other barriers to dispersal disrupting widely distributed species [12, 24-25].

There are also conflicting patterns compared to marine species. Diadromous species should have increased population subdivision due to restriction to freshwater habitats for a portion of their lives. This should increase speciation rates in diadromous species relative to marine species (although some marine reef associated fishes have extremely limited dispersal and a high level of population structure [26-28]). Comparative analyses on a few select taxa suggest that diadromous fishes have intermediate level of population structure compared to the more structured freshwater and less structured marine species [24]. Based on these assumptions, it is expected that diadromous lineages would have intermediate diversification rates if both speciation and extinction decrease compared to freshwater species and increase compared to marine species.

Previous phylogenetic analyses have addressed a second aspect of the evolution of diadromy, the transitions in and out of diadromy. The studies tend to focus on specific families or orders of fishes [2, 19, 22, 30-32]. However, these studies can suffer from low

statistical power owing to small numbers of transitions and little time for speciation and extinction to occur (e.g. Feutry *et al* [19] examine one transition into diadromy and one transition out of diadromy). Some clade-specific studies find no significant difference between marine and freshwater transition rates [23], while other fish groups show considerably more transitions from marine to freshwater compared to the reverse [21]. Some groups have few transitions into diadromy [6, 33], while others have numerous transitions both in and out of diadromy [3,6]. These studies also tend to focus on clades with relatively high prevalence of diadromy. Thus, they might provide insight into the timing and mechanisms of particular transitions but have limited scope for generalization to all fishes. To better assess the evolutionary significance of diadromy, this study uses the most complete and time-calibrated tree of 7,822 fishes constructed using 13 genes and 60 fossils for time calibration [34,35] coupled with MuSSE (Multiple State Speciation and Extinction) [36], a well-established method, to better elucidate the diversification within and transition between three states (i.e. marine, freshwater and diadromous) across ray-finned fishes. The MuSSE model uses a time-calibrated phylogeny to generate estimates of speciation (λ) and extinction (μ) within each state, as well as rates of transition between each state (q_{ij}).

The first aim of this study is to evaluate the importance of diadromy as an intermediate evolutionary stage between marine and freshwater lifestyles (Fig. 2), including a test of the extreme scenario that all transitions between marine and freshwater have to go through diadromy (Fig. 1.2.I). This extreme scenario is contradicted by cases of apparent transition between marine and freshwater without evidence of a diadromous intermediate [21,23]; however, large scale phylogenetic analysis could support the existence of diadromy as a temporary or "hidden" transitional state. More generally, the idea that diadromy is primarily a transitional state would be supported if transitions involving diadromy are quantitatively dominant and diversification of diadromous lineages is relatively unimportant (Fig. 1.2.II). These predictions are compared to a null model with all transition rates being similar regardless of the direction (Fig. 1.2.III).

The second aim of this study is to test the hypothesis that if diadromy only acts as an intermediate state, net diversification rates in diadromous fishes should be lower than either marine or freshwater fishes. Partial movement into new habitats should not result in speciation or extinction, but instead transition out of diadromy. The alternative hypotheses predicts an intermediate diversification rate arising from the expected effects of diadromy on population structure and extinction risk discussed above. It is important to note that both hypotheses are strictly assessing rates and do not address causality in the system. It is not at all clear that alternative causal hypotheses for the evolutionary gain or loss of diadromy predict different *rates* of transition, speciation, or extinction. No attempt is made here to test why or how diadromy evolves. This study focuses on the quantitative importance of diadromy by assessing transition, speciation, and extinction rates of marine, diadromous and freshwater fishes.

Method

Species list:

The 7822 species used in the published phylogeny [29] were assigned one of three character states: 1) diadromous (n=180), 2) freshwater (n=3998), and 3) marine (n=3600). All species that could not be assigned were classified as “unknown” and dropped from the tree (n= 45) (Supplementary file 1.1). Diadromous fishes were identified from 27 families (Salmonidae: n=32, Gobiidae: n=19, Anguillidae: n=18, Clupeidae: n=17, Acipenseridae: n=16, Eleotridae: n=14, Osmeridae: n=9, Galaxiidae: n=7, Cottidae: n=6, Achiridae: n=5, Mugilidae: n=5, Cyprinidae: n=4, Retropinnidae: n=4, Salangidae: n=4, Ariidae: n=2, Gasterosteidae: n=2, Moronidae: n=2, Pleuronectidae: n=2, Rhyacichthidae: n=2, Tetraodontidae: n=2, Cheimarrichthyidae: n=1, Gadidae: n=1, Lateolabracidae: n=1, Latidae: n=1, Plecoglossidae: n=1, Pseudaphritidae: n=1, Syngnathidae: n=1). Character states were determined primarily based on information in: Diadromy in Fishes [6], Fishbase [37], Fishes of the World [38], and Encyclopedia of Life [39]. Species considered as euryhaline, or those that can tolerate a large range of salinity including freshwater, were assigned character states based on their dominant life strategy. *Bairdiella chrysoura*, for example, is listed as marine, freshwater, and brackish in Fishbase, but since the species is described as moving "to the nursery and feeding areas in estuaries during summer months and sometimes enters freshwaters," [37], *B. chrysoura* was listed as a marine species. When inadequate information was available on the aforementioned sources, other primary literature was used in considering salinity levels during breeding, distribution, and salt tolerance, to better assign species to each of the three categories (additional sources used for character assignment: [22, 23, 40-59]). Additionally, the species listed as both marine and freshwater were confirmed via peer-reviewed evidence, such as in the case of *Takifugu poecilonotus*, where studies confirmed diadromy [60]. Species such as *Alosa pseudoharengus* [17], and many of the freshwater eels [61] with both diadromous and landlocked populations, were characterized as diadromous since the behavior is dominant within these species. This was preferable to removing or duplicating these species which would alter measures of speciation and transition between the states.

Phylogenetic model selection:

Species were assigned character states F (freshwater), M (marine), or D (diadromous) for three-state MuSSE analysis. MuSSE gives estimates for speciation (λ), extinction (μ), and transition between states (q_{FD} ; in this example from state F to state D). Net diversification rates (r) were calculated by subtracting extinction rates from speciation rates ($r_i = \lambda_i - \mu_i$). The MuSSE analysis estimates 12 parameters: λ_F , λ_M , λ_D , μ_F , μ_M , μ_D , q_{FM} , q_{FD} , q_{MD} , q_{MF} , q_{DF} , and q_{DM} . Net diversification rate (r) was calculated by subtraction of extinction from speciation ($r = \lambda - \mu$). State Stability (S) was calculated as one minus the probability of extinction of that state minus the probability of transition out of that state to all other states ($S_i = 1 - \mu_i - q_{ij} - q_{ik}$). Similar to other SSE models (e.g. Bisse, HiSSE, GeoSSE), rates (λ , μ and q) are each estimated independently using likelihood starting at the tips of the tree and working back in time along each branch in small time increments.

Estimates for each parameter are compiled at nodes until likelihoods have been calculated to the root. For a full description of parameter estimation see Maddison *et al.* [62], FitzJohn [31] and Ng and Smith [63].

MuSSE analyses were run in R [64] using the “diversitree” package [36]. The correction for incomplete sampling in the tree, the *sampling.f*=(.26,.21,.46) function, was used to inform the program of the percent of freshwater (3,998/15,170) and marine (3,600/16,764) species based on current estimates [65]. Because recognized non-diadromous species occasionally exhibit diadromous behavior, the true number of species that are diadromous remains unknown. The percent of diadromous species used (104/223) was calculated based on the inventory provided in "Diadromy in Fishes," [6] where only 104 of the 223 species listed were present in the phylogeny. A total of 27 models were compared using the MuSSE model (Additional File A1.1). Akaike information criterion (AIC) scores were calculated from the log likelihood value outputs from each model. ΔAIC was calculated by subtracting the lowest AIC value from each AIC score. Akaike weights (w_i) for each model were calculated. Bayesian estimates of the parameters using MCMC, as implemented with the *mcmc* function in DIVERSITREE, were obtained with runs of 10,000 steps. Additionally, to test robustness of character state assignment, models were run with all 45 “unknown” species, which could not be separated in to marine or freshwater, placed completely in marine or completely in freshwater. Results presented in Table 1.1 and Additional file A1.1 do not include these species. All models were run using the single best fit tree reported by Rabosky *et al.* [34] resulting in no variation in tree topology or branch lengths. No additional topologies were made available for analysis. Additional tree topologies and variation in branch lengths could change the estimated rates produced by all models.

Estimates of extinction rates using phylogenies have been suggested to be inaccurate [66], however, studies have shown that in certain situations phylogenies can be good tools for these estimates [67,68]. As pointed out by Beaulieu and O'Meara [68], using a tree with a large number of taxa (>100) should be adequate for rate estimations involving extinction, as long as the variation in birth rate is not exceptionally high (also see Rabosky [69]). To see if varying extinction rates affected other parameters, models were compared with all extinction rates constrained to be equal ($\mu_f \sim \mu_M \sim \mu_D$). Parameter estimates from this model resulted in speciation and transition rates similar to those observed in the best fit models (Table 1.1).

Hidden State Speciation Extinction models (HiSSE) [70] were also run to see if there were any unobservable unmeasured factors (or hidden states) within diadromy. HiSSE is a two-state model that tests for variation or hidden rates within the diadromous and non-diadromous fishes. HiSSE models without hidden states are identical to BiSSE models and were also run. Using the *MarginRecon* function with the best fit model parameters, all diadromous tips were, with >98% likelihood, assigned to only one rate state. As a result, models were run with no hidden state for diadromous fishes. While these models do not allow for testing the hypotheses of whether diadromy is an intermediate state,

identifying a potential hidden state may reveal that a few select taxa are skewing the rates reported in the MuSSE analysis. If there is one particular clade or select species that exhibit a trait that results in a large difference in speciation, extinction or transition rate, it could bias the MuSSE output parameters.

To support the transition rate estimation reported in MuSSE, *fitDiscrete* function in ‘geiger 2.0.3’ [71] was used to estimate transition rates in both a two state (diadromous / nondiadromous) and three state (diadromous / marine / freshwater) data set. Models were compared using AIC. For both two and three-state situations, equal-rate (ER), symmetric (SYM), all rates different (ARD), and meristic (meristic) models were tested.

Model adequacy testing:

To test for model adequacy, trees were simulated with the *tree.musse* function in the diversitree package [36]. Using the best fit MuSSE output parameters, with the ancestral state set as marine, 500 trees were simulated to 7,778 tips. The distribution of the percent of marine, freshwater and diadromous fishes from the 500 simulations were compared to the original percent of taxa in each group. These simulations serve two essential purposes. First, they validate the MuSSE model in that the output parameters properly simulate what is observed in nature. Secondly, they address the potential bias of uneven sampling in the phylogeny used in the study. Showing that results from the tree are representative of what is observed in nature implies that the tree adequately represents the true tree of fishes.

Results

Phylogenetic Hypothesis Testing:

MuSSE results suggest that net diversification rates ($r = \lambda - \mu$) in diadromous fishes ($r = 0.446$) are orders of magnitude higher than in both marine ($r = 0.01$) and freshwater ($r = 0.062$) fishes (Table 1.1, Additional file A1.1). The observed pattern of freshwater fishes having a higher speciation and extinction rate compared to marine fishes is in line with other estimates [32]. The transition rate from diadromy to freshwater ($q_{DF} = 9.37 \times 10^{-3}$) is an order of magnitude higher than transitions in the reverse direction ($q_{FD} = 4.98 \times 10^{-4}$) and the transition rate from diadromy to marine ($q_{DM} = 4.81 \times 10^{-1}$) is two orders of magnitude higher than the reverse ($q_{MD} = 2.87 \times 10^{-3}$). The highest transition rate overall reflected movement from diadromous to marine lifestyles (Table 1.1.A). Results from the best fit model rerun with all “unknown” species placed in marine or freshwater yielded results that maintained similar trends with respect to all output parameters (Additional file A1.2). Posterior distributions show strong variation between diadromous and non-diadromous speciation and extinction rates as well as transition out of diadromy as opposed to other transitions (Fig. 1.3). State Stability (S) was lowest for diadromy ($S_D = 0.51 = 1 - 2.33e^{-7} - 9.37e^{-3} - 4.81e^{-1}$) compared to marine ($S_M = 0.87 = 1 - 1.29e^{-1} - 3.36e^{-4} - 2.87e^{-3}$) and freshwater ($S_F = 0.84 = 1 - 1.55e^{-1} - 1.77e^{-4} - 4.98e^{-4}$)

BiSSE and HiSSE analyses resulted in no support for a hidden state (Additional file A1.3). Due to the nature of these two-state models, comparison between other rates cannot be made, as neither BiSSE nor HiSSE can assess diadromy with respect to its role as a transitional state between marine and freshwater. Instead, the best fit HiSSE model assigns non-diadromous fishes into two rate categories that did not match up with character assignment of marine and freshwater species. Transition rates inferred from fitDiscrete corroborated the MuSSE results showing high transition rates out of diadromy. The ARD model was best fit for both the two state and three state models (Table 1.1.B). Both of the best fit models inferred transition rates out of diadromy that were two orders of magnitude larger than any other transition rate.

Results indicate that diadromy is not necessarily required as an evolutionary transition between marine and freshwater systems, but instead is a rare state that, if obtained, can lead to rapid diversification and large numbers of descendants specialized in either marine or freshwater (Fig. 1.2.IV). If the intermediate state hypotheses [5, 8] were true, we would expect the models constraining transition through diadromy (i.e. $q_{FD} \sim q_{DM}$, $q_{FM} \sim q_{DM}$) or those limiting movement between freshwater and marine (i.e. $q_{FM} \sim 0$ and $q_{MF} \sim 0$) to all be better fit models (Additional file A1.1). In the best-supported model, transitions occur between all states. The top models show highest transition levels while exiting diadromy (Fig. 1.2.IV) and do not show patterns expected from simple interpretations of diadromy as an intermediate state (Fig. 1.2.I, 1.2.III).

Model Adequacy Testing:

Simulations of 500 trees that used the output parameters of the MuSSE model resulted in a distribution of trait frequencies such that the observed values fell within one standard deviation of the mean of the simulated distribution (Fig. 1.4).

Discussion

Substantial transition rates between marine and freshwater in the best fit model do not support the hypothesis that all transitions between marine and freshwater move through diadromy. The patterns observed are neither in support of equal rates across the board (Fig. 1.2.II) nor diadromy as an intermediate with one-to-one transition rates from marine to diadromy to freshwater or the other way around (Fig. 1.2.III). The highest transition rates in both the fitDiscrete and MuSSE analyses were those out of diadromy to both marine and freshwater, which is in line with diadromy having the lowest stability ($S_D = 0.51$). Additionally, transitions from marine to diadromy were observed at an intermediate rate and those from freshwater to diadromy and between marine and freshwater showed the lowest rates. Many mechanisms could be responsible for the high rates of transition exiting diadromy including landlocked and oceanlocked populations. Freshwater populations can become isolated each time a set of random individuals becomes landlocked due to factors such as: a receding glacier forming isolated lakes, change in flow patterns of rivers, or the formation of natural dams due to tectonic processes. Oceanlocked species can result from a loss of suitable freshwater habitat in a

species whose populations are geographically limited to ephemeral island streams. These high transition rates out of diadromy are illustrated by many examples of diadromous species with non-diadromous individuals, populations, or sister species [6, 18, 17, 72]. Species in which ocean- and landlocking has resulted in genetic isolation include alewives, galaxiids, gobies, sticklebacks, and lampreys [6, 17, 73]. With multiple ocean- or landlocking events and a wide enough distribution, multiple completely-freshwater species, or completely-marine species in the case of eels [18, 74], can emerge from a single diadromous species. Once species have diverged, fine scale niche preferences can support a diversified clade in their newly occupied habitats [75]. The directionality of observed transitions suggest that diadromy may sometimes be an intermediate state between marine and freshwater life histories and sometimes a temporary state in which lineages persist before returning to their ancestral state. Many papers address the causes and mechanisms of these transitions [5,8,9,19]; however, those are beyond the scope of this study.

The high net-diversification rate of diadromous fishes is somewhat unexpected due to the potential for dispersal (especially in species without a strong tendency to return to their natal sites), since these fishes are capable of movement between distant freshwater sources during their marine phase. A wide range of dispersal distances have been reported in diadromous fishes [25,67,76]. The variation in dispersal time and distance is important in speciation and extinction rates. Short dispersal potential can result in a high level of isolation while still allowing for occasional propagules to venture out and form new populations with low levels of gene flow, a classic scenario for geographic speciation [77]. Long distance dispersal can result in less structured populations compared to poorly dispersing diadromous fishes but could increase the overall range of the species allowing for more opportunities for isolation and transitions out of diadromy, especially in species that exhibit facultative diadromy. These results contradict expectations that diadromy has an intermediate diversification rate associated with an intermediate level of gene flow and range size compared to marine and freshwater fishes. Instead, diadromous fishes have higher diversification rates than either freshwater or marine lineages. The estimated speciation rate is about twice that of freshwater species and four times that of marine species (Table 1.1). This might arise if diadromous species, rather than having intermediate population structure, experience the best of both worlds for speciation. That is, perhaps diadromous species tend to have large geographic ranges (more like marine species on average), but also high levels of subdivision and local adaptation (more like freshwater species on average). Even more striking, the estimated extinction rate is several orders of magnitude lower for diadromous vs non-diadromous lineages (Table 1.1). Taken together with the high transition rates, it appears that diadromous ancestors might often escape extinction by leaving relictual freshwater populations in the case of unfavorable conditions in the ocean, and likewise by persisting in the ocean when freshwater populations are extirpated.

While these results neither support or refute the proposed hypothesis with respect to the causality of the evolution of diadromy, they do not conflict with the proposed transitions

patterns proposed by Gross [8], Feutry [19], and Dodson [9]. These results support a hypothesis that a high rate of transitions out of diadromy coupled with the high net diversification rate in diadromy leads to a single species becoming diadromous, diversifying and then many species returning to non-diadromous lives. Using the flagtails as an example, Feutry *et al.* [19] shows that they have a marine ancestry followed by a transition into diadromy. Once diadromy evolved, multiple speciation events occurred giving rise to six extant diadromous species of *Kuhlia*. Then, presumably via the gradual transition of the partially diadromous *K. munda*, a transition back to completely marine occurred. Of the diadromous species half of them (i.e. *K. malo*, *K. rupestris* and *K. munda*) exhibit intermediate phenotypic variation and/or partially catadromous life histories. While extant species show only one well defined transition out of diadromy, the loss of total obligate diadromy in some of the diadromous species could be additional transition out of diadromy.

In Salmonids, there are two major transitions into diadromy and many transitions out of diadromy. Alexandrou *et al.* [78] show a transition from freshwater to diadromy and back to freshwater within Coregoninae. Within the Coregoninae clade the transition into diadromy has resulted in eight diadromous species. Either *Coregonus artedi* was a second transition into diadromy after a transition back to freshwater, or transition from diadromy to freshwater occurred both in *C. nigripinnis* and in the *C. zenithicus* and *C. hoyi* clade. However inconsistencies in the placement of various taxa in *Coregonus* [79; 80] could drastically change the order of events with regards to transitions and any assumptions would be speculative. If we are to assume the Crete-Lafreniere *et al.* [79] phylogeny is correct, two transitions out of diadromy are observed. First, from the widespread *C. autumnalis*, the Ireland endemic *C. pollan* arises. Second, within the polyphyletic *C. artedi* is nested a group of freshwater US Great Lakes species. In both cases a transition from diadromy to freshwater is occurring in a temperate region, opposite to what Gross proposed [8]. Within the species rich Salmoninae clade, Alexandrou *et al.* [78] show an additional two transitions out of diadromy, in the genera *Salmo* and *Salvelinus*. In all topologies, we see many diadromous species arise and at least one, if not multiple transitions to freshwater, a pattern that coincides with best fit MuSSE predictions regardless of the selective forces that have been imposed on these systems.

Within the Anguilliformes, one major transition from marine to freshwater has occurred in the freshwater eels of the genus *Anguilla* resulting in the 16 species and three subspecies [81]. In agreement with the prediction of Gross *et al.* [81], we find marine ancestors transitioning to a catadromous life history in a tropical region. There is evidence that in tropical regions individuals of *Anguilla marmorata* [83] and in temperate regions *A. japonica* and *A. Anguilla* in temperate regions [74] have reverted back to completely marine life histories. While these directions do not always fit the hypotheses proposed by Gross *et al.* (1988) and Dodson *et al.* (2009), each can be thought of as another potential transition out of diadromous life history.

Many landlocking events have been observed in the threespine stickleback (*Gasterosteus aculeatus*), for which multiple mechanisms have been proposed [16, 84]. Despite the genetic and life-history variation between diadromous and freshwater *G. aculeatus* populations, hybridization does occur [16], though no studies report gene flow between distinct, geographically isolated freshwater populations (although introgressive hybridization between sympatric benthic-limnetic species pairs has occurred [85]). It is possible that with current population dynamics, each landlocked population could eventually lead to distinct taxa, since adaptive radiation has allowed morphological and behavioral variation to accrue in landlocked populations within a short time. In the case of the alewife (*Alosa pseudoharengus*), contact between many landlocked populations and the diadromous population does not occur, and genetic evidence indicates that many of the landlocked populations are the result of separate isolation events such as damming of a river or glacial retreat [17]. Additionally, there are morphological, behavioral and genetic differences between landlocked and diadromous populations that have diverged within the last century [86]. Studies have even started to pinpoint specific functional genes that may play a role in the ability to survive landlocking events [86]. Though no speciation has been recognized in *A. pseudoharengus*, these independent landlocking events are the result of multiple transitions out of diadromy [17].

Conclusion

Complex life histories that require environmental transitions can serve as much more than evolutionary stepping stones between different life histories. Diadromy provide species with the ability to enter new environments and diversify even if it is a behavior that may be quickly lost. Applying SSE models to other complex life histories may reveal more patterns of biodiversity. Additionally, considering complex life histories that require environmental transitions not only in terms of their costly transitions, but for the advantages of using multiple environments, may allow for new perspectives on multiple topics related to life histories and evolution.

List of abbreviations

AIC: Akaike information criterion, **ARD**: all rates different, **BiSSE**: Binary State Speciation and Extinction, **ER**: equal-rate, **HiSSE**: Hidden State Speciation and Extinction, **MuSSE**: Multiple State Speciation and Extinction, **SYM**: symmetric

Acknowledgements

I would like to thank Jeremy Beaulieu, Brian O'Meara, and Orlando Schwery for their help in SSE model analyses, Darrin Hulsey and Brent Tibbits for their discussions on diadromous fishes, and Jeremy Beaulieu, Graciela Cabana, Ben Fitzpatrick, Ben Keck, Brian O'Meara, and Danial Simberloff for their comments throughout the project as well as two anonymous reviewers for their comments. K. Lindsay Hunter copy edited sections

of the manuscript. Funding was provided by the University of Tennessee Department of Ecology and Evolutionary Biology.

References

1. Shubin, N. H., E. B. Daeschler, and F. A. Jenkins. The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature*. 2006;440(7085):764-771.
2. Dudley, R. et al. Gliding and the functional origins of flight: biomechanical novelty or necessity? *Annu Rev Ecol Evol S*. 2007;38:179-201.
3. Bloom, D. D. and N. R. Lovejoy. The evolutionary origins of diadromy inferred from a time-calibrated phylogeny for Clupeiformes (herring and allies). *P Roy Soc B-Biol Sci*. 2014;281(1778):2013-2081.
4. Myers, G. S. Usage of anadromous, catadromous and allied terms for migratory fishes. *Copeia*. 1949;2:89-97.
5. Gross, M. R. Evolution of diadromy in fishes. In *Am Fish S S*. 1987;1:14-25.
6. McDowall, R. M. *Diadromy in Fishes: Migrations Between Freshwater and Marine Environments*. Timber Press, Portland, OR. 1988.
7. Hodges, M. H. and F. W. Allendorf. Population genetics and pattern of larval dispersal of the endemic Hawaiian freshwater amphidromous gastropod *Neritina granosa* (Prosobranchia: Neritidae). *Pac Sci*. 1998;52(3):237-249.
8. Gross, M. R., R. M. Coleman, and R. M. McDowall. Aquatic productivity and the evolution of diadromous fish migration. *Science* 1988;239(4845):1291-1293.
9. Dodson, J. J., J. Laroche, and F. Lecomte. Contrasting evolutionary pathways of anadromy in euteleostean fishes. In *Challenges for diadromous fishes in a dynamic global environment*. *Am Fish S S*. 2009;69:63-77.
10. McCormick, S. D., P. F. Anthony, and J. B. Colin. eds. *Fish Physiology: Euryhaline*, 32 Academic Press. Waltham MA. 2013.
11. Black, V.S. Excretion and osmoregulation. *The physiology of fishes*. 1957;(1):163-205
12. Castelin, M. et al. New insight on population genetic connectivity of widespread amphidromous prawn *Macrobrachium lar* (Fabricius, 1798) (Crustacea: Decapoda: Palaemonidae). *Mar Biol*. 2013;160(6):1395-1406.
13. Kuribayashi, K., H. Katakura, M. Kyono, M. H. Dick, and S. F. Mawatari. Round-trip catadromous migration in a Japanese amphipod, *Sternomoera rhyaca* (Gammaridea: Eusiridae). *Zool Sci*. 2006;23(9):763-774.
14. Chen, L. H, J. L. Yang. and M. D. Barnes. Effect of natural disturbance on migrations and population dynamics of a catadromous species, *Eriocheir japonicus*, along the Lo-Mei stream in northern Taiwan. *Hydrobiologia*. 2014;726(1):179-194.
15. Parenti LR. Life history patterns and biogeography: an interpretation of diadromy in fishes. *Annals of the Missouri Botanical Garden*. 2008;95(2):232-47.
16. McPhail, J. D. Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of south-western British Columbia p.399-437. In *The Evolutionary Biology of the Threespine Stickleback*. 1994.
17. Palkovacs, E. P., B. D. Kirstin, D. M. Post, and A. Caccone. Independent evolutionary origins of landlocked alewife populations and rapid parallel evolution of phenotypic traits. *Mol Ecol*. 2008;17(2), 582-597.

18. Tsukamoto, K. and N. Izumi. Do all freshwater eels migrate? *Nature*. 1998;396(6712):635-636.
19. Feutry, P. et al. Evolution of diadromy in fish: insights from a tropical genus (*Kuhlia* species). *Am Nat*. 2013;181(1):52-63.
20. Lee, C. E., and M. A. Bell. Causes and consequences of recent freshwater invasions by saltwater animals. *Trends Ecol Evol*. 1999;14(7):284-288.
21. Betancur-R, R. Molecular phylogenetics supports multiple evolutionary transitions from marine to freshwater habitats in ariid catfishes. *Mol. Phylogenet Evol*. 2010;(55):249-258.
22. Betancur-R, R., Ortí, G., & Pyron, R. A. Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecol Lett*. 2015;18(5):441-450.
23. Bloom, D.D. and Lovejoy, N.R. On the origins of marine-derived freshwater fishes in South America. *J Biogeogr*. 2017;44(9):1927-1938.
24. McDowall, R. M. Diadromy, diversity and divergence: implications for speciation processes in fishes. *Fish Fish*, 2001;2(3):278-285.
25. Shiao, J. C., C.S. Tzeng, P. C. Li, and K. N. Bell. 2015. Upstream migration and marine early life history of amphidromous gobies inferred from otolith increments and microchemistry. *Environ Biol Fish*, 98(3), 933-950.
26. D'Aloia, C.C., Bogdanowicz, S.M., Harrison, R.G. and Buston, P.M., 2014. Seascape continuity plays an important role in determining patterns of spatial genetic structure in a coral reef fish. *Molecular Ecology*, 23(12), pp.2902-2913.
27. Alfaro, M.E., Santini, F. and Brock, C.D., 2007. Do reefs drive diversification in marine teleosts? Evidence from the pufferfish and their allies (Order Tetraodontiformes). *Evolution: International Journal of Organic Evolution*, 61(9), pp.2104-2126.
28. Taylor, M.S. and Hellberg, M.E., 2006. Comparative phylogeography in a genus of coral reef fishes: biogeographic and genetic concordance in the Caribbean. *Molecular Ecology*, 15(3), pp.695-707.
29. Ward, R.D., Woodward, M. and Skibinski, D.O.F. A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *J Fish Biol*. 1994;44(2):213-232.
30. McDowall, R. M. The evolution of diadromy in fishes (revisited) and its place in phylogenetic analysis. *Rev Fish Biol. Fisher* 1997;7(4):443-462.
31. Betancur-R, R., Ortí, G., Stein, A. M., Marceniuk, A. P., & Alexander Pyron, R. Apparent signal of competition limiting diversification after ecological transitions from marine to freshwater habitats. *Ecol Lett*. 2012;15(8):822-830.
32. Bloom, D. D., J. T. Weir, K. R. Piller, and N. R. Lovejoy. Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of New World silversides (Atherinopsidae) *Evolution*. 2013;67(7):2040-2057.
33. Watanabe, K., Sakai, H., Sanada, T. and Nishida, M., 2018. Comparative phylogeography of diadromous and freshwater daces of the genus *Tribolodon* (Cyprinidae). *Ichthyological Research*, pp.1-15.

34. Rabosky, D. L., Santini, F., Eastman, J., Smith, S. A., Sidlauskas, B., Chang, J. *et al.* Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat Commun.* 2013a;4: 1958.
35. Rabosky, D. L., Santini, F., Eastman, J., Smith, S. A., Sidlauskas, B., Chang, J. *et al.* Data from: Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Dryad Digital Repository. [doi:10.5061/dryad.j4802](https://doi.org/10.5061/dryad.j4802). 2013b
36. FitzJohn, R. G. Diversitree: comparative phylogenetic analyses of diversification in R. *Method Ecol Evol.* 2012;3(6):1084-1092.
37. Froese, R. and D. Pauly. Editors. FishBase. World Wide Web electronic publication. www.fishbase.org 2014.
38. Nelson, J. S. *Fishes of the World.* John Wiley and Sons). 2006.
39. Wilson, E. O. The encyclopedia of life. *Trends Ecol Evol.* 2003;18:77-80.
40. Frost, W. E. The natural history of the minnow, *Phoxinus phoxinus*. *J Anim Ecol.* 1943;12:139-162.
41. Breder, C. M. *Modes of Reproduction in Fishes.* Natural History Press. Garden City, NJ. 1966.
42. Foster, N. R. Behavior, development, and early life history of the Asian needlefish, *Xenentodon cancila*. *P Acad Nat Sci Phila.* 1973;125:77-88.
43. Govoni, J. J. and J. V. Merriner. The occurrence of ladyfish, *Elops saurus*, larvae in low salinity waters and another record for Chesapeake Bay. *Estuaries.* 1978;1(3):205-206.
44. Linden, E., B. E. Bengtsson, O. Svanberg, G. Sundström. and N. S. E. P. Board. The acute toxicity of 78 chemicals and pesticide formulations against two brackish water organisms, the bleak (*Alburnus alburnus*) and the harpacticoid *Nitocra spinipes*. *Chemosphere.* 1979;8(11-12):843-851.
45. Allen, G. R. *Snappers of the World: An Annotated and Illustrated Catalogue of Lutjanid Species Known to Date.* FAO Fisheries Synopsis, no. 125, vol. 6. Food and Agriculture Organization. Rome. 1985.
46. Goto, A. Alternative life-history styles of Japanese freshwater sculpins revisited. *Environ Bio Fish.* 1990;28:101-112.
47. Marais, J. F. K. and D. J. L. Venter. Changes in body composition associated with growth and reproduction in *Galeichthys feliceps* (Teleostei: Ariidae). *S Afr J Marine Sci.* 1991;10(1):149-157.
48. Rosecchi, E. and A. J. Crivelli. Study of a sand smelt (*Atherina boyeri* Risso 1810) population reproducing in fresh water. *Ecol of freshw fish.* 1992;1(2):77-85.
49. Sakai, H. Life-histories and genetic divergence in three species of *Tribolodon* (Cyprinidae). *Memoirs of the Faculty of Fisheries, Hokkaido University,* 1995;(1-2):1-98.
50. Rainboth, W. J. *Fishes of the Cambodian Mekong.* (Food and Agriculture Org). 1996.
51. Garg, S. K. Brackishwater carp culture in potentially waterlogged areas using animal wastes as pond fertilizers. *Aquacult Int.* 1996;4(2):143-155.

52. Bemis, W. E., E. K. Findeis, and L. Grande. An overview of Acipenseriformes. *Environ Biol Fishes*. 1997;48(1-4), 25–71.
53. Natsumeda, T. Year-round local movements of the Japanese fluvial sculpin, *Cottus pollux* (large egg type), with special reference to the distribution of spawning nests. *Ichthyol Res*. 1999;46(1):43-48.
54. Takeshita, N., N. Onikura, S. Nagata, S. Matsui, and S. Kimura. A note on the reproductive ecology of the catadromous fourspine sculpin, *Cottus kazika* (Scorpaeniformes: Cottidae). *Ichthyol Res*. 1999;46(3):309-313.
55. Lovejoy, N. R. and B. B. Collette. Phylogenetic relationships of New World needlefishes (Teleostei: Belontiidae) and the biogeography of transitions between marine and freshwater habitats. *Copeia*. 2001;2:324-338.
56. Chen, Y. F. Induced ovulation and embryonic development of ocellated puffer, *Takifugu ocellatus*. *J Appl Ichthyol*. 2005;21(2):136-140.
57. Toyama, M., *et al*. Presence of Rhodopsin and Porphyropsin in the Eyes of 164 Fishes, Representing Marine, Diadromous, Coastal and Freshwater Species—A Qualitative and Comparative Study. *Photochem and photobiol*. 2008;84(4):996-1002.
58. Hua, X. *et al*. Phylogeographical analysis of an estuarine fish, *Salanx ariakensis* (Osmeridae: Salanginae) in the north-western Pacific. *J Fish Biol*. 2009;75(2):354-367.
59. Machado, I., M. Vera, D. Calliari, and L. Rodríguez-Graña. First record of an *Elops smithi* (Pisces: Elopidae) larva in a South American subtropical-temperate estuary. *Mar Biodivers Records*. 2012;5:e108.
60. Kato, A., H. Doi, T. Nakada, H. Sakai, and S. Hirose. *Takifugu obscurus* is a euryhaline fugu species very close to *Takifugu rubripes* and suitable for studying osmoregulation. *BMC physiology*. 2005;5(1):18.
61. Thibault, I., J. J. Dodson, F. Caron, W. N. Tzeng, Y. Iizuka, and J. C. Shiao. Facultative catadromy in American eels: testing the conditional strategy hypothesis. *Mar Ecol-Prog Ser*. 2007;344:219-229.
62. Maddison, W.P., Midford, P.E. and Otto, S.P., 2007. Estimating a binary character's effect on speciation and extinction. *Systematic biology*, 56(5), pp.701-710.
63. Ng, J. and Smith, S.D., 2014. How traits shape trees: new approaches for detecting character state-dependent lineage diversification. *Journal of evolutionary biology*, 27(10), pp.2035-2045.
64. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2013.
65. Eschmeyer, W. N., R. Fricke, J. D. Fong, and D. A. Polack. Marine fish diversity: history of knowledge and discovery (Pisces). *Zootaxa*. 2010;2525(1):19-50.
66. Rabosky, D. L. Extinction rates should not be estimated from molecular phylogenies. *Evolution*. 2010;64(6), 1816-1824.
67. Nee, S., E. C. Holmes, R. M. May, and P. H. Harvey. Extinction rates can be estimated from molecular phylogenies. *Philos T Roy Soc B*. 1994;344(1307):77-82.
68. Beaulieu, J. M. and B. C. O'Meara. Extinction can be estimated from moderately sized molecular phylogenies. *Evolution*. 2015;69(4):1036-1043.

69. Rabosky, D. L. Challenges in the estimation of extinction from molecular phylogenies: A response to Beaulieu and O'Meara. *Evolution*. 2016;70(1), 218-228.
70. Beaulieu, J.M. and O'Meara, B.C. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst Biol*, 2016;65(4):583-601
71. Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, W. Challenger. GEIGER: investigating evolutionary radiations, *Bioinformatics*. 2008;24(1):129-131
72. Augspurger, J.M., M. Warburton, and G. P. Closs. Life-history plasticity in amphidromous and catadromous fishes: a continuum of strategies. *Rev Fish Biol Fisher*. 2017;27:1-16.
73. Bracken, F. S., A. R. Hoelzel, J. B. Hume, and M. C. Lucas. Contrasting population genetic structure among freshwater-resident and anadromous lampreys: the role of demographic history, differential dispersal and anthropogenic barriers to movement. *Mol Ecol*. 2015;24(6):1188-1204.
74. Daverat, F. et al. Phenotypic plasticity of habitat use by three temperate eel species, *Anguilla anguilla*, *A. japonica* and *A. rostrata*. *Mar Ecol-Prog Ser*. 2006;308:231-241.
75. Glova, G. J., Jellyman, D. J. & Bonnett, M. L. Factors associated with the distribution and habitat of eels (*Anguilla* spp.) in three New Zealand lowland streams. *NZ J Mar Freshw Res*. 1998;32:255-269.
76. Feutry P, Valade P, Ovenden JR, Lopez PJ, Keith P. Pelagic larval duration of two diadromous species of Kuhliidae (Teleostei: Percoidei) from Indo-Pacific insular systems. *Mar Freshwater Res*. 2012;63(5):397-402.
77. Coyne, J.A. and Orr, H.A., 2004. Speciation. Sunderland, MA.
78. Alexandrou, M.A., Swartz, B.A., Matzke, N.J. and Oakley, T.H., 2013. Genome duplication and multiple evolutionary origins of complex migratory behavior in Salmonidae. *Molecular phylogenetics and evolution*, 69(3), pp.514-523.
79. Crête-Lafrenière, A., Weir, L.K. and Bernatchez, L., 2012. Framing the Salmonidae family phylogenetic portrait: a more complete picture from increased taxon sampling. *PLoS One*, 7(10), p.e46662.
80. Horreo, J.L., 2017. Revisiting the mitogenomic phylogeny of Salmoninae: new insights thanks to recent sequencing advances. *PeerJ*, 5, p.e3828.
81. Inoue, J.G., Miya, M., Miller, M.J., Sado, T., Hanel, R., Hatooka, K., Aoyama, J., Minegishi, Y., Nishida, M. and Tsukamoto, K., 2010. Deep-ocean origin of the freshwater eels. *Biology Letters*, 6(3), pp.363-366.
82. Arai, T. and Chino, N., 2018. Opportunistic migration and habitat use of the giant mottled eel *Anguilla marmorata* (Teleostei: Elopomorpha). *Scientific reports*, 8(1), p.5666.
83. Bell, M. A. and C. A. Andrews. Evolutionary consequences of colonization of fresh water by primitively anadromous fishes. Pp. 323-363 in B. Streit, T. Stadler and C. M. Lively eds. *Evolutionary ecology of freshwater animals. Current Trends and Case Studies*, Birkhauser, Basel, Switzerland. 1997.
84. Czesny, S., J. Epifanio. and P. Michalak. Genetic divergence between freshwater and marine morphs of alewife (*Alosa pseudoharengus*): a 'next-generation' sequencing analysis. *PloS one* 2012;7(3),e31803.

85. Velema, G.J., J.S. Rosenfeld, and E.B. Taylor. Effects of invasive American signal crayfish (*Pacifastacus leniusculus*) on the reproductive behaviour of threespine stickleback (*Gasterosteus aculeatus*) sympatric species pairs. *Can J Zool*, 2012: 90(11):1328-1338.
86. Michalak, K., Czesny, S., Epifanio, J., Snyder, R. J., Schultz, E. T., Velotta, J. P. et al. Beta-thymosin gene polymorphism associated with freshwater invasiveness of alewife (*Alosa pseudoharengus*). *J Exp Zool Part A*. 2014;321(4):233-240.

Appendix A – Figures and Tables

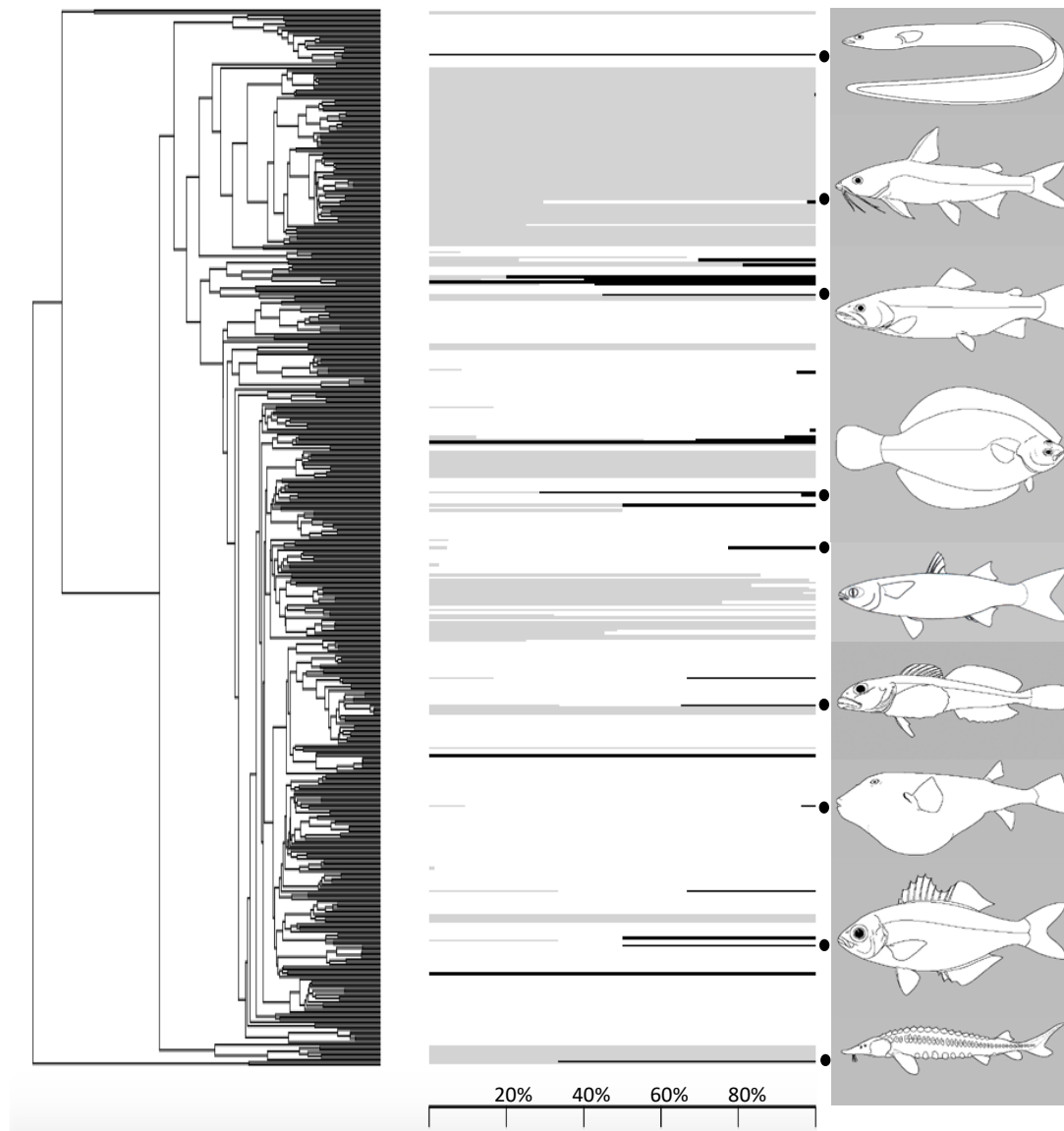


Figure 1.1 Relative distribution of diadromy in fishes.

Phylogeny from Rabosky *et al.* (2013b) pruned to the family level. Horizontal lines represent the percent of freshwater (gray), marine (white), and diadromous (black) species from each family represented in the original tree. Note that percentages are based only on species used in the original phylogeny. For example, in the family Pseudaphritidae two species are recognized, one diadromous and one marine. This study only incorporates *Pseudaphritis urvillii*, the diadromous species, so the horizontal line for the family Pseudaphritidae is completely black despite the only 50% of the family being diadromous. Each family is represented by only one tip and placement of non-

monophyletic families was chosen based on the placement of the majority of its species. Dots next to pictures (30) represent placement of select families as a general reference. Families represented are, from top to bottom, Anguillidae, Ariidae, Salmonidae, Pleuronectidae, Mugilidae, Cottidae, Tetraodontidae, Kuhliidae, Acipenseridae.

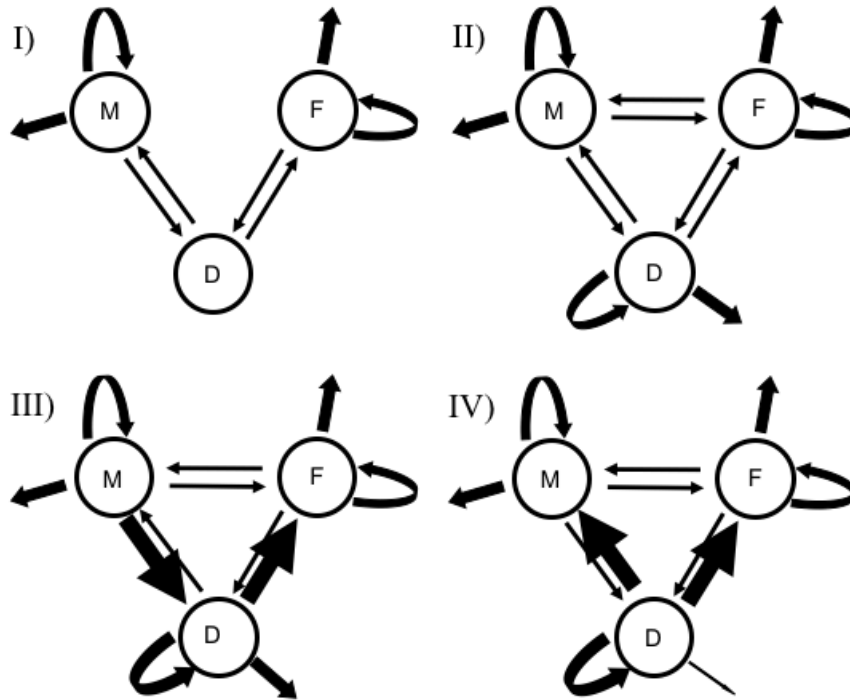


Figure 1.2 Illustrated hypotheses of the evolution of diadromy. Circles represent marine (M), freshwater (F) and diadromous states (D). Arrows between states represent transitions (q_{ij}), arrows leaving and returning to the same state represent speciation rate (λ_i), and arrows leading away from each state represent extinction (μ_i). Size of arrows within each category (speciation, extinction and transitions) are representative of the rate variation, but not sized to scale. I) Model hypothesizing that diadromy is a transitional state between marine and freshwater. This hypothesis should be better supported if diadromy prevails only as an evolutionary stepping stone to other character states. II) Hypothetical model where diadromy is evolutionarily stable and its prevalence is influenced by variation in rates of speciation and extinction. III) Hypothesis showing movement between marine and freshwater utilizing diadromy as a stepping stone, while still allowing for direct movement between marine and freshwater. IV) Optimal models based on MuSSE results.

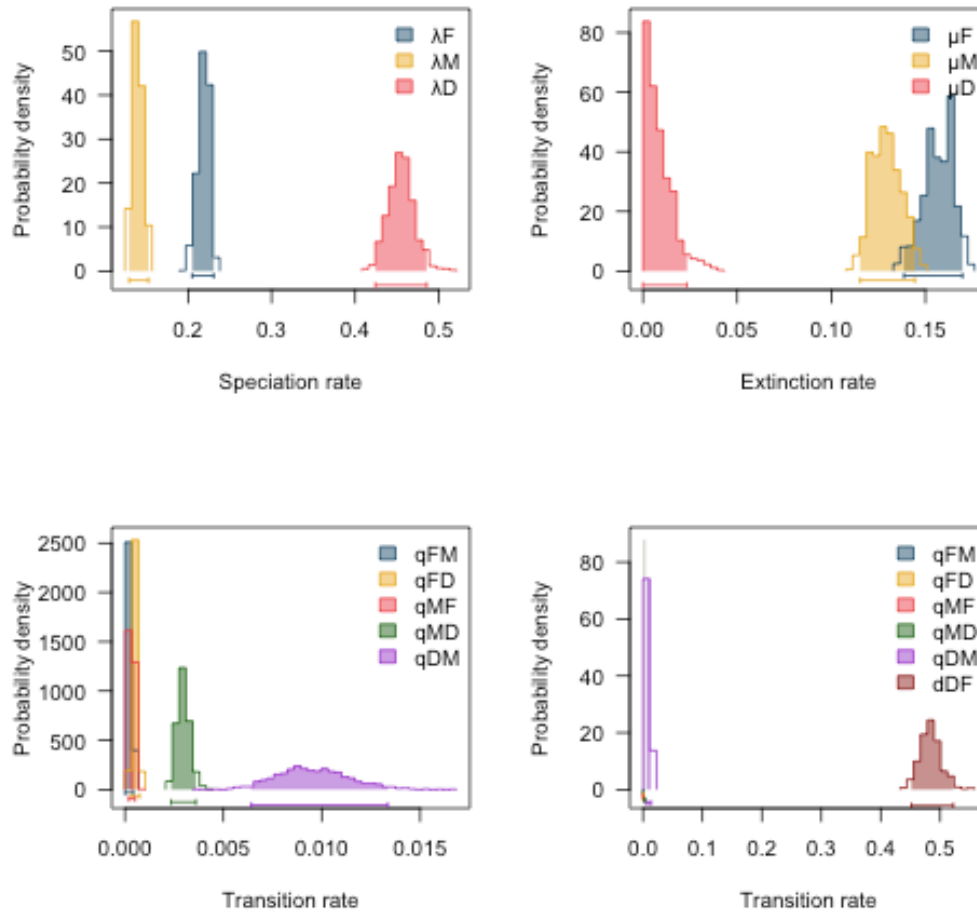


Figure 1.3 Posterior probability distributions for MuSSE output paramaters. Distributions shown for all output paramaters including speciation (λ_F , λ_M , and λ_D), extinction (μ_F , μ_M , and μ_D), and transition (q_{FM} , q_{FD} , q_{MD} , q_{MF} , q_{DF} , and q_{DM}) rates inferred by the MuSSE. Distributions are based on 10,000 generation runs with a burn-in of the 10%.

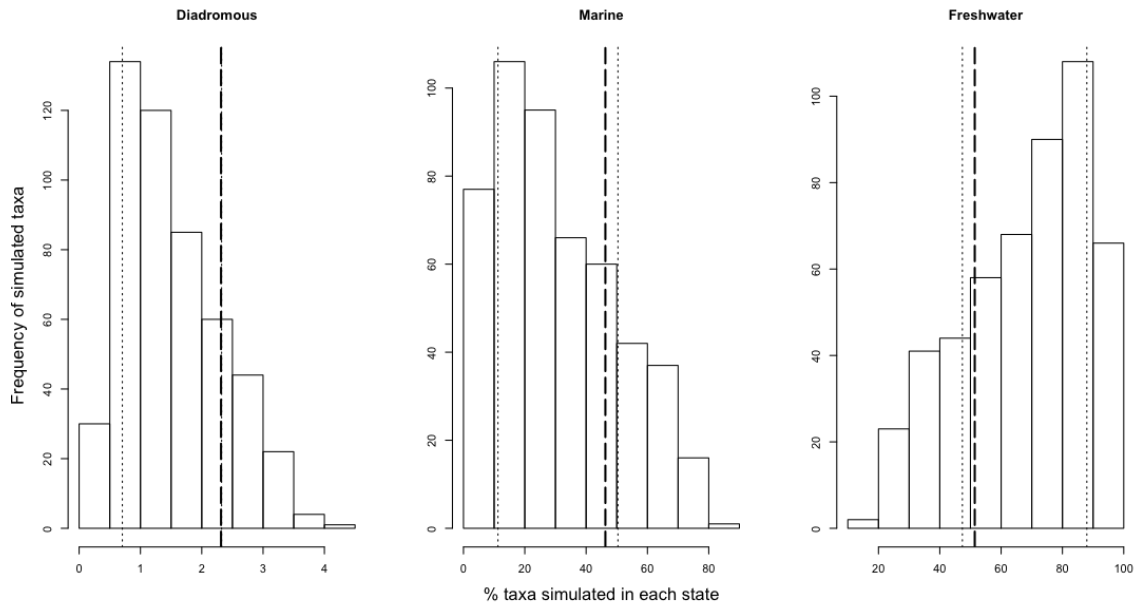


Figure 1.4 Simulated distribution of marine, freshwater, and diadromous fishes. Histograms of percent taxa in each state (freshwater, marine, and diadromous) from 500 simulated trees using the output parameters from the optimal MuSSE model. All trees were assigned marine as ancestral state. Thick dashed lines indicate the observed values for each state. Thin dotted lines indicate one standard deviation from the mean of simulated values.

Table 1.1

Output parameters for key model in MuSSE and fitDiscrete analyses.

A)

Full (No constraints)		Δ AIC: 0	wi: 0.380		
λ_F	λ_M	λ_D	μ_F	μ_M	μ_D
2.17E-01	1.39E-01	4.46E-01	1.55E-01	1.29E-01	2.33E-07
q_{FM}	q_{FD}	q_{MF}	q_{MD}	q_{DF}	q_{DM}
1.77E-04	4.98E-04	3.36E-04	2.87E-03	9.37E-03	
qFM ~ qFD		Δ AIC: 0.09	wi: 0.363		
λ_F	λ_M	λ_D	μ_F	μ_M	μ_D
2.17E-01	1.39E-01	4.45E-01	1.55E-01	1.29E-01	1.46E-07
q_{FM}	q_{FD}	q_{MF}	q_{MD}	q_{DF}	q_{DM}
3.82E-04	3.28E-04	3.38E-04	2.91E-03	9.57E-03	4.79E-01
qFM ~ 0		Δ AIC: 0.79	wi: 0.257		
λ_F	λ_M	λ_D	μ_F	μ_M	μ_D
2.16E-01	1.38E-01	4.48E-01	1.54E-01	1.28E-01	7.58E-06
q_{FM}	q_{FD}	q_{MF}	q_{MD}	q_{DF}	q_{DM}
0	5.95E-04	3.23E-04	2.88E-03	9.28E-03	4.84E-01
$\mu_F \sim \mu_M \sim \mu_D$		Δ AIC: 27.45	wi: 4.16E-07		
λ_F	λ_M	λ_D	μ_F	μ_M	μ_D
2.04E-01	1.48E-01	5.36E-01	1.39E-01	1.39E-01	1.39E-01
q_{FM}	q_{FD}	q_{MF}	q_{MD}	q_{DF}	q_{DM}
1.88E-04	6.28E-04	3.07E-04	3.12E-03	7.20E-03	4.37E-01

B)

ARD 3-state		Δ AIC: 0			
q_{FM}	q_{FD}	q_{MF}	q_{MD}	q_{DF}	q_{DM}
3.55E-04	1.64E-04	3.82E-04	4.44E-04	1.75E-02	1.19E-02
ARD 2-state		Δ AIC: 0			
q_{ND}			q_{DN}		
1.88E-04			2.85E-02		

Output parameters for A) MuSSE best fit model and model with constraints on extinction rates. Speciation rate (λ_i), extinction rate (μ_i), and transition rates (q_{ij}) where i and j refer to the original and new states respectively. B) fitDiscrete models for both two and three state models.

Appendix B - Additional file

Table A1.1

Model constraints for all MuSSE analyses.

MuSSE constraints	Δ AIC	wi
Full (No constraints)	0	0.38
$q_{FM} \sim q_{FD}$	0.09	0.363
$q_{FM} \sim 0$	0.79	0.256
$\mu_F \sim \mu_M, \mu_D \sim \mu_M$	27.45	4.16E-07
$q_{FD} \sim 0$	84.41	1.78E-19
$q_{FM} \sim 0, \mu_F \sim \mu_M, \mu_D \sim \mu_M$	133.66	3.60E-30
$\mu_F \sim \mu_M, \mu_D \sim \mu_M, q_{FM} \sim q_{FD}$	205.05	1.13E-45
$q_{MF} \sim q_{MD}$	206.66	5.06E-46
$q_{DF} \sim q_{DM}$	207.27	3.73E-46
$q_{FM} \sim q_{MF}$	208.56	1.96E-46
$q_{MF} \sim 0$	238.35	6.65E-53
$\mu_F \sim \mu_M, \mu_D \sim \mu_M, q_{FM} \sim q_{FD}, q_{MF} \sim q_{MD}, q_{DF} \sim q_{DM}$	245.31	2.05E-54
$q_{MF} \sim 0, q_{FM} \sim 0$	265.34	9.17E-59
$q_{FM} \sim q_{DM}$	279.5	7.72E-62
$q_{MD} \sim q_{DM}$	287.69	1.29E-63
$q_{FM} \sim q_{FD}, q_{DM} \sim q_{FD}$	307.48	6.48E-68
$q_{DM} \sim 0$	309.31	2.60E-68
$q_{MD} \sim 0$	321.33	6.37E-71
$q_{FD} \sim 0, q_{DM} \sim 0$	324.81	1.12E-71
$q_{MF} \sim q_{DF}$	421.66	1.04E-92
$q_{FD} \sim q_{DF}$	421.69	1.03E-92
$\mu_F \sim \mu_M, \mu_D \sim \mu_M, \lambda_F \sim \lambda_M, \lambda_D \sim \lambda_M$	428.28	3.80E-94
$q_{FM} \sim 0, q_{FD} \sim q_{DM}$	483.08	4.79E-106
$q_{DF} \sim 0$	569.69	7.47E-125
$q_{MD} \sim 0, q_{DM} \sim 0$	605.94	1.00E-132
$q_{MD} \sim 0, q_{DF} \sim 0$	612.51	3.76E-134
$q_{FD} \sim 0, q_{DF} \sim 0$	815.65	2.910E-178

Model constraints for MuSSE analyses. Constraints involving speciation rate (λ_i), extinction rate (μ_i), and transition rates (q_{ij}) where i and j refer to the original and new states respectively. Akaike weights (wi) were calculated for each model using equation (1). $\sim$ indicates that a rate was set equal to a second rate or to 0.

Table A1.2

Output parameters for optimal model in MuSSE analyses with “unknown” species in either freshwater or marine.

Unknown Taxa in Freshwater					
λ_F	λ_M	λ_D	μ_F	μ_M	μ_D
2.15E-01	1.40E-01	4.51E-01	1.54E-01	1.29E-01	1.39E-06
q_{FM}	q_{FD}	q_{MF}	q_{MD}	q_{DF}	q_{DM}
3.51E-04	4.79E-04	3.51E-04	2.81E-03	1.17E-02	4.83E-01

Unknown Taxa in Marine					
λ_F	λ_M	λ_D	μ_F	μ_M	μ_D
2.18E-01	1.37E-01	4.51E-01	1.56E-01	1.26E-01	5.01E-07
q_{FM}	q_{FD}	q_{MF}	q_{MD}	q_{DF}	q_{DM}
3.44E-04	4.29E-04	3.44E-04	2.90E-03	1.08E-02	4.86E-01

Speciation rate (λ_i), extinction rate (μ_i), and transition rates (q_{ij}) where i and j refer to the original and new states respectively.

Table A1.3
Models constraints for HiSSE and BiSSE analysis.

A)

HiSSE constraints	Δ AIC	wi
pp27 ta6, ea6, tm8	0	1
<i>pp30 ta6, ea6, tm9</i>	100.45	1.54014E-22
pp5: ta4, ea5, tm2	231.54	5.269E-51
<i>pp28: ta7, ea7, tm8</i>	2878.2	0
<i>pp62: ta4, ea4, tm2</i>	1929.48	0
<i>pp31: ta7, ea7, tm9</i>	2970.82	0
pp32: ta8, ea8, tm9	3033.57	0
<i>pp19: ta2, ea2, tm6</i>	3103.28	0
<i>pp22: ta1, ea1, tm10</i>	3103.28	0
pp8: ta3, ea2, tm6	3107.01	0
pp29: ta8, ea8, tm8	3111.55	0
pp20: ta3, ea3, tm6	3112.23	0
pp2: ta3, ea3, tm7	3309.32	0
<i>pp9: ta2, ea3, tm7</i>	3322.46	0
<i>pp7: ta3, ea2, tm7</i>	3390.88	0
<i>pp1: ta2, ea2, tm7</i>	3398.48	0
<i>pp10: ta2, ea3, tm6</i>	3643.07	0
<i>pp: ta1, ea1, tm1</i>	2.00E+10	0
<i>pp4: ta5, ea5, tm2</i>	2.00E+10	0
<i>pp6: ta4, ea5, tm12</i>	2.00E+10	0

B)

net turnover	extinction fraction
ta1=(1,1,1,1)	ea1=(1,1,1,1)
ta2=(1,1,0,0)	ea2=(1,1,0,0)
ta3=(1,2,0,0)	ea3=(1,2,0,0)
ta4=(1,2,3,4)	ea4=(1,2,3,4)
ta5=(1,2,1,2)	ea5=(1,2,1,2)
ta6=(1,2,3,0)	ea6=(1,2,3,0)
ta7=(1,2,1,0)	ea7=(1,2,1,0)
ta8=(1,1,1,0)	ea8=(1,1,1,0)

A1.3 cont.

transition rate matrix									
<u>tm1</u>	(0A)	(1A)	(0B)	(1B)	<u>tm8</u>	(0A)	(1A)	(0B)	(1B)
(0A)	NA	4	7	10	(0A)	NA	3	5	0
(1A)	1	NA	8	11	(1A)	1	NA	6	0
(0B)	2	5	NA	12	(0B)	2	4	NA	0
(1B)	3	6	9	NA	(1B)	0	0	0	NA
<u>tm2</u>	(0A)	(1A)	(0B)	(1B)	<u>tm9</u>	(0A)	(1A)	(0B)	(1B)
(0A)	NA	3	5	0	(0A)	NA	3	4	0
(1A)	1	NA	0	7	(1A)	1	NA	0	0
(0B)	2	0	NA	8	(0B)	2	0	NA	0
(1B)	0	4	6	NA	(1B)	0	0	0	NA
<u>tm6</u>	0	1			<u>tm10</u>	(0A)	(1A)	(0B)	(1B)
0	NA	2			(0A)	NA	1	4	0
1	1	NA			(1A)	1	NA	0	6
					(0B)	2	0	NA	7
<u>tm7</u>	0	1			(1B)	0	3	5	NA
0	NA	1							
1	1	NA							

A) Models with ΔAIC and model weights for HiSSE and BiSSE models. B) Net turnover (ta), Extinction fractions (ea), and transitions matrices (tm) used in models. For Net turnover and Extinction fractions list represent (0A, 1A, 0B, 1B) where 0 and 1 are non-diadromous and diadromous and A and B represent observed and hidden rates. ta2, ta3, ea2, ea3, tm6 and tm7 are used in models with no hidden states. Ta6, ta7, ta8, ea6, ea7, ea8, tm8, and ta9 are used in models where a hidden state is observed in non-diadromous, but not diadromous fishes.

Appendix C -Attachments

Attachment 1.1

Species character state assignment.

SuppTable1.xls

CHAPTER II
GENETIC DIVERSITY AND POPULATION STRUCTURE OF
MARINE, CATADROMOUS, ANADROMOUS, AMPHIDROMOUS
AND FRESHWATER FISHES

Co-authors on this chapter include Evin T. Carter and Benjamin M. Fitzpatrick. The Chapter was written by JBC. JBC, ETC and BMF designed the study, JBC collected the data. JBC and ETC conducted the spatial analyses, JBC analyzed the data. All co-authors assisted with revisions and writing the manuscript.

Abstract

Phylogenetic comparative methods have shown that diadromous fishes, or those with an obligate scheduled movement between marine and freshwater, have a high rate of diversification compared to completely marine and freshwater taxa. A high propensity for speciation leads to the prediction that diadromous species will often have disjunct populations with restricted gene flow and a strong pattern of isolation by distance. Although previous studies have concluded that diadromous species have intermediate levels of genetic diversity (marine species tend to be higher and freshwater species lower heterozygosity at allozyme loci), the partitioning of variation between populations has not been compared. We compared patterns of isolation by distance for 31 species representing marine and freshwater specialists, and all forms of diadromy (anadromous, catadromous and amphidromous). We additionally compared expected heterozygosity (H_e) across 82 species to evaluate whether microsatellite data support earlier conclusions based on allozymes. Our results do not support an association between migration life history and population genetic structure. We find no significant differences between life histories in either population differentiation or within population heterozygosity - in contrast to previous conclusions. Upon review, those previous conclusions are based on weak statistical comparisons. Our results suggest that differentiation in populations may not be as strongly associated with diversification rates as previously thought. Advanced methods in understanding geographic movement are critical for understanding population structure.

Introduction

Genetic diversity within and between populations is a function of geographic range and life history (Wright 1943; McCulloch *et al.* 2009; Harvey *et al.* 2017a). The dynamics of population geographic subdivision are key to understanding speciation (Mayr 1963, Coyne and Orr 2004, Bolnick and Fitzpatrick 2007). It is expected that rates of speciation should be related to levels of population variation within a species (Harvey *et al.* 2017b, but see: Kisel *et al.* 2012). Isolation by distance (IBD) should, all else being equal, cause more widespread species to have high rates of speciation (Wright 1943). However, many factors, including movement (dispersal and migration), reproductive output, body size, lifespan, range size, local abundance and genetic variation all interact in populations and alter the theoretical results proposed by Wright (1943). Body size in particular is strongly associated with metabolism, range, reproductive output, genetic variation and diversification in lineages (Hutchinson and MacArthur 1959; Peters 1983; Kochmer and Wagner 1988). Differences in migratory life history have been associated with variation in diversification rates (Corush Ch1) and should therefore also be associated with patterns of population structure. Here, we address the hypothesis that body size and migratory life history jointly predict population genetic structure.

Life history and genetic differentiation between populations

Population connectivity is altered by barriers and movement of individuals. Changes in life history that affect these components should affect the population structure of a species. Adaptations like flight, ability to go onto land or into water and the ability to move between fresh and saltwater can all change how individuals move across their species' range. If a freshwater fish can move into the ocean and back into a different watershed (Bracken *et al.* 2015), if a completely aquatic insect can leave the water and fly to another lake to reproduce (McCulloch *et al.* 2009), or if lice undergo phoretic dispersal as opposed to self-dispersal (Dibiasi *et al.* 2018), then the resulting gene flow should reduce the population subdivision of that species. Conversely, when those abilities are lost, the resulting barriers should decrease gene flow and increase population subdivision (McDowall 2001).

In fishes, marine and freshwater systems present different physical barriers, which have different levels of permeability for fish. Marine systems tend to have less well defined and more permeable barriers because they are vast bodies of continuous water where organisms are expected to have high levels of population connectivity (Waples 1987; Palumbi 1994). Variation in depth, sea surface temperature, and vicinity to shelter, such as coral reefs or kelp forests, do affect the propensity for a fish to move from one location to another (Schultz *et al.* 2008; Selkoe 2010), but there are species that freely move across the globe each generation (Ely *et al.* 2005). Larger, more continuous distributions should decrease population subdivision. While some fishes have a strong preference for specific habitats that may affect their movement (e.g., sponge gobies [D'Aloia *et al.* 2013]), in many species the physical barriers are less clear-cut. The effects of this factor

on population subdivision may contribute to the lower rate of speciation found in marine lineages (Bloom and Lovejoy 2013, Corush Ch1).

Freshwater taxa are restricted by saltwater and land, in addition to thermal and habitat gradients. Small island streams, smaller river drainages on large continents and isolated lakes all limit gene flow to a smaller geographic area. Free movement of even a few hundred kilometers in many freshwater systems is hampered by waterfalls, dams, ephemeral streams and land or salt water. The continuous ranges of freshwater fishes in most examples, even in the case of the Amazon and Nile Rivers, are small in comparison to ranges in the oceans. All things being equal, these habitat restrictions are expected to result in lower population sizes. Together, these factors may contribute to high levels of speciation and endemism found in freshwater fishes (Smith 1992; Etnier and Starnes 1993; Miller *et al.* 2005; Bloom and Lovejoy 2013, Corush Ch1).

The ability to breach the barriers imposed on freshwater taxa would greatly change the population dynamics and restrictions imposed on populations. Diadromous fishes, or those that exhibit a scheduled movement between marine and freshwater, can regularly move between isolated freshwater systems (Myers 1949). Three major variants of diadromy are anadromy, catadromy and amphidromy (Myers 1949; Gross 1987; McDowall 1988). Anadromous fishes reproduce in freshwater and migrate to marine systems for growth and development (e.g. salmon). Catadromous fishes reproduce in marine systems and migrate to freshwater for growth and development (e.g., freshwater eels). Amphidromous fishes reproduce in freshwater, migrate to the sea for a portion of their development (usually <200 days), and then migrate back to freshwater to continue growing prior to reproduction (e.g., Hawaiian waterfall climbing goby). Unlike euryhaline fishes, diadromous species are generally not free to move across salinity gradients at all life stages.

Diadromous behavior allows a species that occurs in freshwater habitats to cross the saltwater boundary imposed on most rivers. Movement into new different freshwater habitat after the marine phase increases gene flow across the species' range. It also allows for the potential for a species to move into previously uncolonized waterways thereby increasing the species' range and the potential number of individuals. With a large geographic expansion, and the increased gene flow, the increased opportunity for isolation by distance and local adaptation should increase the species' population subdivision and genetic diversity (McDowall 2001). Each form of diadromy is affected by this habitat permeability in different way.

Anadromous species are similar to freshwater taxa in that reproduction takes place in an isolated area. Even if there are resources for a large number of individuals to live in the ocean, they are limited by habitat availability during reproduction. Species with strong homing, such as some salmon, go back to their natal stream for reproduction (Garant *et al.* 2000) and would be expected to display population genetic patterns similar to freshwater species. Although species with strong homing may wander during the marine

phase (Leider 1989), they are expected to show high levels of population subdivision. Species with lower or no homing may move between streams at different rates, increasing gene flow, lowering the level of population subdivision.

Amphidromous species have a limited time in marine systems as the larvae feed and grow at sea. The time spent in the ocean is typically less than in anadromous species (McDowall 1988). However, amphidromous species are known for their dispersal to remote islands (McDowall 2004, 2007). In some small systems such as species endemic to the Hawaiian Islands, amphidromy allows for admixture across a species range (Chubb *et al.* 1998). However, the more consistent use of freshwater should limit population size and dispersal time, causing amphidromous population structure to be most similar to that of freshwater species.

Catadromous species are different from the other forms of diadromy because they are less constrained during their reproductive stages. Because they can use many freshwater habitats, restrictions in resources in rivers should not affect the overall species. For example, the American eel (*Anquilla rostrata*), which spawns in a very specific site in the Sargasso Sea, uses freshwater habitats from Mexico to Greenland (Schmidt 1923). This species has few physical restrictions on gene flow across its breeding range. Catadromous fishes should be more similar to marine fishes than freshwater species.

Life history and genetic diversity within populations

Total genetic diversity (H_e) in a species is a function of number of individuals and movement between populations. Freshwater species typically have lower population sizes, less movement between populations, and should have less gene diversity within populations compared to marine fishes that have large population sizes, increased movement between populations, and higher gene diversity within populations (Ward *et al.* 1994; DeWoody and Avise 2000). Diadromous fishes' population dynamics increase movement between populations and number of individuals compared to completely freshwater fishes, but have an increased set of restrictions compared to marine fishes (McDowall 2001). Anadromous, amphidromous, and catadromous fishes should fall between the high diversity of marine and low diversity of freshwater fishes. Even with migration out of freshwater, factors such as homing should decrease the amount of migrations between populations and preserve the freshwater population dynamics. However, there are ample examples of straying in fishes that have strong homing tendencies, which can have a large effect on within population variation (Quinn and Fresh 1984; Keefer and Caudill 2014).

Population genetics of diadromy

The differences in population structure caused by variation in migration life history should be reflected in the genetic diversity found in various species (Gyllenstein 1985; Ward *et al.* 1994; DeWoody and Avise 2000; McDowall 2001). Freshwater resident lampreys show greater between population variation compared to sympatric anadromous lampreys (Bracken *et al.* 2015). In the galaxiids of New Zealand, diadromous populations

had lower genetic variation between sites compared to completely freshwater populations (Allibone and Wallis 1992). Comparisons between marine and freshwater stickleback population showed higher levels of heterozygosity in marine populations (Jones *et al* 2012). With more structured populations, it is expected that fishes' recruitment should occur at a smaller spatial scale. As expected, freshwater fishes have a smaller recruitment area compared with an intermediate recruitment area in anadromous and a large recruitment area in marine fishes (~500km: marine > anadromous > ~50km: freshwater) (Myers *et al.* 1997). Gyllensten (1985), Ward *et al.* (1994), and DeWoody and Avise (2000) reported differences in genetic diversity across species with different life histories. Gyllensten (1985) assessed allozymes (primarily from salmonids) and concluded that mean total gene diversity of marine fishes ($H_t = 0.063$) is higher than freshwater ($H_t = 0.043$) and anadromous ($H_t = 0.041$) fishes. Ward *et al.* (1994) evaluated a greater number and a more taxonomically diverse set of examples and again used allozyme data to come to the similar conclusion that the gene diversity averaged across populations decreased from marine to anadromous to freshwater. However, they found no significant difference between groups with respect to total genetic variation. DeWoody and Avise (2000) then used microsatellites to show a statistically significant level of variation in heterozygosity between marine ($H = 0.77$) and freshwater species ($H = 0.54$) with anadromous species ($H = 0.68$) falling in between. Higher average within-population genetic diversity of marine vs. freshwater fishes was attributed to higher effective population size (hence maintenance of neutral genetic variation). Although not statistically distinguishable from either extreme, the intermediate diversity of anadromous populations was interpreted as consistent with intermediate effective population sizes (DeWoody and Avise 2000). However, these previous analyses ignored the possibility of phylogenetic non-independence and did not consider other life history traits associated with genetic variation as a potential confounding variable associated with effective population size.

Many aspects of a species biology are important in population structure, including body size (Olden *et al.* 2007; Luiz *et al.* 2011; Strona *et al.* 2012). Characteristics such as maximum length have been associated with range size (Strona *et al.* 2012), which affects potential for increased population subdivision or isolation by distance (Wright 1943). Diadromous species tend to grow larger than their freshwater counterparts (Hardisty and Potter 1971; Thériault *et al.* 2007). Diadromous species also tend to have high reproductive output compared to freshwater species (Closs *et al.* 2013). All these life history traits are interconnected, and each plays a role in population structure.

As compared to previous studies, this study adds a series of additional components for a more thorough comparison of diadromy and its relevance in population dynamics. First, we incorporate additional life histories (i.e. catadromy and amphidromy) to understand the complex variation in diadromous migrations. We incorporate additional traits that are known to be associated with genetic variation (e.g. body size). We test for a phylogenetic signal. Lastly, we conduct a special analysis using genetic vs. geographic distance comparisons to reassess the effects of diadromy on population genetics in fishes. We use published molecular studies to assess the effects of life history on the level of population

structure as reflected in the relationship between genetic and geographic distance. We collected sampling locations to calculate pairwise geographic distances and the associated pairwise F_{st} values between each pair of populations. Additionally, we collected published heterozygosity (H_e) data from species in each life history to re-evaluate the previous conclusion that marine species have higher within-population variation than freshwater species, with diadromous species intermediate. Evolutionary relatedness was corrected for in all analyses. We predict that 1) H_e will be lowest in freshwater and increase sequentially in amphidromous, anadromous, catadromous and marine fishes and 2) marine fishes will have the lowest levels of population structure because of their large population sizes and wide geographic range. Catadromous fishes will have a slightly increased level of population structure because of their restriction to freshwater. Anadromous species will have an even greater level of population structure due to both the restrictions to movement caused by their freshwater phase in addition to the restrictions imposed during breeding. Amphidromous species will have a high level of population structure because the majority of their life and their breeding takes place in freshwater. Freshwater species should have the highest level of population structure.

Methods

Phylogeny

To correct for phylogenetic non-independence, a phylogeny was estimated using multiple published trees as suggested by Beaulieu *et al.* (2012). A phylogeny was produced using PhyloT (<http://phylot.biobyte.de/>) including the following taxa: *Oreochromis niloticus*, *Myxine*, *Negaprion brevirostris*, *Cyprinus carpio*, *Danio rerio*, *Takifugu*, *Mola*, *Camptostoma*, *Percina*, *Umbra*, *Kuhlia*, *Periophthalmus*, *Pyura*, *Ichthyomyzon*, *Petromyzon*, *Eptatretus*, *Dasyatis*, *Squalus*. All teleost tips were dropped and replaced with a pruned version from the Rabosky *et al.* (2013) phylogeny. If particular species were not in the tree, a sister taxon was chosen that would not affect branch lengths. For example, *Agonostomus telfairii* was not in the Rabosky phylogeny, but *Agonostomus monticola* was, and since no other fish in the monophyletic group containing those species was used in this study, switching names has no effect on any branch lengths or topology in the current comparison. Tips and node dates for non-teleosts were manually adjusted based on dates in the literature as follows: Most recent common ancestor (MRCA) Orectolobiformes and Carcharhiniformes: ~289 Million years ago (MYA), MRCA *Galeocerdo cuvier* and *Negaprion brevirostris*: ~72 MYA, MRCA cartilaginous and bony fish: ~525(494–580) MYA, MRCA Gnathostomes and Agnathans: ~652 (605–742) MYA (Kumar and Hedges 1998; Blair and Hedges 2005; and Melo *et al.* 2016). All tree manipulation was done using Ape (Paradis *et al.* 2004) in R (R core team) (Figure A2.1 and A2.2).

To test for associations between life history and within and between population genetic diversity (See below), all phylogenetic linear regressions were run using the inferred tree and the `phylolm()` function in the `phylolm` package (Ho and Ane 2014). All phylogenetic linear regressions models were set to “lambda” in order to estimate Pagel’s λ , which is a

transformation of branch lengths (Ranging from 0-1 where 1 indicates no transformation of branch lengths and 0 reverting to a star tree). All analyses were also run using `phylolm()` with a star tree as well as `lm()`, which does not incorporate any phylogenetic information.

Life history and genetic differentiation between populations

For pairwise F_{st} comparisons, a minimum of four populations (six pairwise comparisons), 10 individuals, and five loci must have been reported. All pairwise F_{st} values were calculated using Weir and Cockerham (1984). Population locations were compiled from each study. If Global Positioning System (GPS) coordinates were not provided, locations were identified based on maps or location descriptions reported in each study (Table A2.1).

For each pair of locations in the 31 data sets containing pairwise F_{st} , two measurements of geographic distance were estimated. The distance between two points as a straight line (Euclidian distance) and a shortest path distance (SPD) (shortest distance between two points via water). GPS coordinates were projected on a map using WGS84 projection in ArcMap and Euclidian distance between each pair of populations was calculated. Multiple surface water layers (Data from Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, IGN, and AeroGRID) were compiled and merged together using the merge tool to create one global surface water shape file. Each river containing a data point was manually inspected to ensure connectivity to the ocean and other river systems was not lost as a function of non-overlapping layers (i.e. a river layer stops right before it goes into the ocean layer which would eliminate that as a potential pathway). Any spurious breaks in waterways were manually added after referencing detailed maps of the area. SPD was then estimated using the created water layer.

To assess isolation by distance, a measure of genetic variation between populations ($F_{ST}/1-F_{ST}$) was regressed against both measures of geographic distance (Rousset 1997). For each species the slope of a) genetic variation vs. Euclidian distance and b) genetic variation vs. SPD was used as response variable in a phylogenetic linear regression (`phylolm()`) with life history and log transformed maximum length as predictor variables. All models were run using the inferred tree, a star tree and no phylogeny (Table 2.2). Mantel and partial mantel tests were then used to assess the correlation between genetic and geographic distance for each species. For each analysis 1000 permutations were used (Table 2.3). Additionally, t-tests with Bonferroni correction for multiple comparisons were used to identify any significant differentiation between pairs of life histories.

Life history and genetic diversity within populations

Studies using microsatellites on marine, anadromous, catadromous, amphidromous, and freshwater fishes were collected. Expected within-population heterozygosity (H_e) was collected from each study. At least 10 individuals and at least 3 loci must have been reported from one or more populations. Maximum length as reported on fishbase (Froese

and Pauly 2018) was collected for each species (Table 2.1). Log transformation was applied to body size due to the non-normal distribution of the data. No additional data (e.g. weight, length at maturity, age at maturity, reproductive output) were available for all species.

Phylogenetic Linear Regression was used to assess heterozygosity as a function of life history and log (maximum length). All models were run with the inferred tree, star tree and no phylogeny (Figure A2.2). Additionally, t-tests with Bonferroni correction for multiple comparisons were used to identify any significant differentiation between pairs of life histories.

Results

Life history and genetic differentiation between populations

Euclidian distance and shortest path distance were calculated between each population and slopes of ($F_{ST} / 1 - F_{ST}$) and geographic distance were calculated for each species (Table 2.3, Table A2.5). Pairwise F_{ST} as a function of geographic distance was free of any phylogenetic signal ($\lambda = 2.69E-09$ for best fit model) (Table 2.2). All models containing shortest path distance were better fit (higher r^2) compared to the same data with Euclidian distance. Two of the five slopes estimated for freshwater species were negative, and none were statistically distinguishable from zero (Table 2.3). Marine, anadromous, and catadromous species had appreciable numbers of significant positive slopes.

No statistical support was found for any model explaining variation in isolation by distance among species (Table 2.2). No models were supported over the null. Although there appears to be a weak trend toward steeper IBD slopes in anadromous compared to catadromous species ($t=2.3585$, $df=11.0$, $p=0.0379$), this might be an artifact of differences in average body size (Table 2.2). However, there was no significant variation found between any two life histories' slopes after correction for multiple comparisons using the Bonferroni correction

Life history and genetic diversity within populations

Heterozygosity within sample sites (H_e) of fishes ranged from $H_e = 0.022$ (catadromous *Macquaria colonorum*) to $H_e = 0.917$ (marine *Syngnathus scovelli*) with an overall mean of 0.68. Mean heterozygosities for each life history were: freshwater: 0.653, amphidromous: 0.633, anadromous: 0.694, catadromous: 0.661, and marine: 0.719. Maximum length ranged from 3.4cm (*Rhinogobius rubromaculatus*) to 750.0 cm (*Galeocerdo cuvier*) (Table 2.1, Figure 2.1). No significant variation was found between any two life histories' heterozygosity after implementing the Bonferroni correction for multiple comparisons. No support was given to any linear regression using life history and/or Log(max length) to explain heterozygosity.

Mean maximum length for each life history was: freshwater: 74.7cm, amphidromous: 30.3, anadromous: 57.4, catadromous: 96.3, and marine: 132.4 (Table 2.1 Figure

2.2). While there was a trend in variation in maximum length between amphidromous and marine ($t=2.836$, $df=25.52$, $p=0.008$), amphidromous and anadromous ($t=2.175$, $df=12.3$, $p=0.0498$), and amphidromous and catadromous ($t=2.917$, $df=9.55$, $p=0.016$), no significant variation was found between any two life histories' maximum length after the Bonferroni correction for multiple comparisons were implemented.

Discussion

Life history and genetic diversity between populations

No models tested showed significant association between IBD and life history and/or maximum body length. While anadromous, catadromous, and marine species showed the expected trend (less isolation by distance in catadromous and marine species), freshwater and amphidromous species showed no support for any IBD (Table 2.3). Although the number of case studies is small, this result indicates that factors such as physical barriers, historical stream capture events, and habitat choice often overwhelm the effect of distance on genetic differentiation for species depending largely or entirely on freshwater. It may be the case that other measurements of geographic distance may need to be applied for a more realistic understanding of movement. Levels of IBD showed a large range of variation especially in marine, catadromous and amphidromous species. The numerous negative slopes may indicate that the way geographic measurements of distance are made between aquatic organisms need to incorporate more landscape features that are biologically relevant to specific species. SPD is one of the few that applies to all fishes and is measurable in a comparative framework. Some species, such as *Anguilla anguilla*, *Sicyopterus stimpsoni*, *Lamoptera planeri*, *Siphateles bicolor*, and *Leucopsarion petersii*, showed drastic changes in their IBD slopes when one goes from Euclidian to SPD. Even SPD assumes that the fish are swimming in straight lines from one population to another, which is known to be inaccurate (Armsworth 2001; Green and Fisher 2004; Béguyer-Pon *et al* 2015). Incorporation of other variables such as currents, water temperature, depth, and so on may give a better approximation of the distance the fish are actually traveling.

There is a lack of support for the notion that the high rate of speciation found in diadromy and in freshwater compared to marine fishes (Bloom and Lovejoy 2013; Corush Ch1) is reflected in the population genetics of those species. This suggests that population differentiation might not be directly linked to diversification. Many lineages may become distinct as a function of isolated populations but become evolutionary dead ends or rejoin their parent lineage before becoming distinct entities resulting in high levels of population differentiation and low levels of diversification. Conversely, quick isolating events may drive rapid speciation before metapopulation dynamics are observable, resulting in low levels of population differentiation and high levels of diversification. But also, high levels of genetic variation may not result in speciation. Fishes also have a high rate of hybridization (Hubbs 1955) suggesting reproductive isolation in a species might require more than increased genetic differentiation. Conversely introgression of another species' genome from a hybridization event can cause increased genetic diversity.

Disassociation between these two processes has been observed in other taxa as well. In orchids of Costa Rica there was no observable pattern in the population variation in species from sister clades that show high and low levels of diversification (Kisel *et al.* 2012). In related examples, birds and *Drosophila* show no observed connection between intrinsic reproductive isolation and diversification rates (Rabosky and Matute 2013). There are, however, many examples that do support the notion that population variation and diversification are linked (Harvey *et al.* 2017b).

Life history and genetic diversity within populations

If species spending all or most of their lives in the ocean tend to have larger effective population sizes than those more dependent on freshwater, then we expect to see greater gene diversity in marine and anadromous species. Previous expectation and compilations of allozyme and microsatellite data were consistent with this expectation (Gyllensten 1985; Ward *et al.* 1994; DeWoody and Avise 2000; McDowall 2001). However, our analysis indicates that the pattern is weak or non-existent in microsatellite data. The observed trends in the mean heterozygosity in each life history do follow those previously observed but lack statistical support. Each life history showed a wide range of heterozygosity. Although smaller body size is also expected to be associated with larger effective population size, we found no significant association between maximum length and heterozygosity. Incorporation of a larger number of studies of heterozygosity could remove a lot of the noise of the data resulting in less variation and a significant pattern

Assessing IBD and H_e in a comparative framework.

Heterozygosity and IBD can change drastically from species to species. A lack of overall trend may also be a function of the scale at which the comparison is being made. Heterozygosity is not consistent or conserved over larger clades or even between closely related species (Switzer *et al.* 2008; Neilson and Stepien 2011, Bracken *et al.* 2015) although there are many examples where heterozygosity does remain fairly similar (Krieg *et al.* 1999; Palkovacs *et al.* 2014; Feldheim *et al.* 2014). The high heterozygosity in the freshwater *Neogobius pallasii* ($H_e=0.77$) might actually be low when compared to its marine relatives, and the low heterozygosity of *Eucyclogobius newberryi* ($H_e=0.313$), might be high compared to its freshwater relatives. Unfortunately, there are too few genetic studies to do a large-scale sister group comparison across all five life histories. The same concept should apply to maximum length.

Maximum length seems to be most associated to within and between population genetic variation. This, however, does not necessarily suggest that there is no impact of life history on the genetics of a species. In comparisons of salmonids, anadromous individuals tend to grow larger than freshwater residents (Thériault *et al.* 2007). Anadromous lampreys tend to get larger than the closely related freshwater lampreys (Hardisty and Potter 1971). Other life history traits such as reproductive output are greater in diadromous compared to non-diadromous species (McDowall 1988; McDowall 1970; Closs *et al.* 2013). In the threespine stickleback, anadromous females have a larger clutch size and bigger bodies than freshwater females (Baker 1994). A species range size

is connected to the reported max length, which is affected by diadromy. However, these observed patterns in size might be lost in a large-scale comparative analysis and should be addressed in a sister group comparison instead.

Multiple caveats are in order here owing to methodological variation among studies. The studies that were aggregated contain a large amount of variations: Variation in the life history stage at which sampling occurs may affect the observable pattern of population structure. Many of the studies of both catadromous and anadromous fishes collect samples during their freshwater stages. This is appropriate for anadromous species, but not necessarily for catadromous species if freshwater sites do not correspond to particular marine breeding sites. If catadromous species were sampled at multiple marine breeding sites, they might show some population structure. Variation in markers used may be an additional source of discontinuity between compared studies. Di- tri- and tetranucleotide repeats each have different mutation rates (Chakraborty *et al.* 1997; Schug *et al.* 1998) and thus more variation would be expected in studies that have more di-nucleotide repeats. A relatively small number of species limits the power of a comparative study. With limited population genetics studies utilizing consistent marker types in fishes, in particular diadromous species, we see a wide spread in the data. This is particularly apparent in freshwater and amphidromous species. These species seem to be the most complex in their population structuring. Information about the natural history, geographic history and evolutionary history of these species needs to be incorporated to obtain a better understanding of how life history is affecting a species' population structure.

Available natural history information is also limited and difficult to incorporate across species. Here we were able to obtain maximum length for each species, however, reproductive output, age of maturity, size at maturity and percent survivorship to maturity are all pieces of information that are also critical in understanding gene diversity, yet often not available. These data, along with mutation rates, can be used to get estimates of how much genetic variation we should expect, allowing for a more nuanced comparison between life histories. To better understand IBD, information about larval duration, passive vs. active larval swimming behavior, feeding behaviors in larvae, adult movement analyses and fine scale niche preferences could be used to see if habitat use is really structuring populations. For example, preference in lentic compared to lotic waters could change how two species in the same river move within and between populations. Feeding behavior, size and larval dispersal distance all result in different patterns of gene flow between populations and can in different way affect IBD (Strathmann 1985; Palumbi 2003). Association between larval duration and extinction risks suggests that larval duration can influence levels of isolation across populations (Douglas *et al.* 2013).

Geographic information relating to physical breaks in their environments caused by waterfalls, ephemeral streams, disconnection between populations and water currents all add levels of complexity that cannot be easily assessed in a comparative framework (Koizumi *et al.* 2006). Historical events can also leave long-lasting signals on populations. For example, changes in sea levels can isolate freshwater habitat such

Taiwan and mainland Asia (Corush, Ch3). Glaciation in the northern hemisphere has drastically shaped the connectivity of freshwater systems, potentially disrupting patterns of isolation by distance expected based on equilibrium assumptions (Wright 1943).

Evolutionary history of a species is also important in assessing IBD. Variation in rate or required changes (Molecular, physiological or behavioral) in reproductive isolation can cause some groups to become distinct entities with low levels of population differentiation, whereas other groups with slower rates of reproductive isolation may show patterns of greater IBD as a result of the ability to maintain unity despite higher levels of population differentiation. Examples in salmonids suggest that even with low genetic variation species can become reproductively isolated (Ryman *et al.* 1979; Hendry *et al.* 2000). This would result in the high level of IBD being overlooked by the high speciation rate.

Together these factors add multiple layers of complexity to understand IBD in aquatic systems. It seems that simple measurements of distance, across a representative number of samples and populations does not adequately allow for the assessment of how life history affects a fish's microevolutionary patterns.

Acknowledgments:

We would like to thank Graciela Cabana, Ben Keck, and Danial Simberloff for their comments throughout the project.

References

- Adams, S.M., Lindmeier, J.B. and Duvernell, D.D., 2006. Microsatellite analysis of the phylogeography, Pleistocene history and secondary contact hypotheses for the killifish, *Fundulus heteroclitus*. *Molecular Ecology*, 15(4), pp.1109-1123.
- Arigoni, S. and Largiader, C.R., 2000. Isolation and characterization of microsatellite loci from the ocellated wrasse *Symphodus ocellatus* (Perciformes: Labridae) and their applicability to related taxa. *Molecular Ecology*, 9(12), pp.2166-2169.
- Armsworth, P.R., 2001. Directed motion in the sea: efficient swimming by reef fish larvae. *Journal of theoretical biology*, 210(1), pp.81-91.
- Beaulieu, J.M., Ree, R.H., Cavender-Bares, J., Weiblen, G.D. and Donoghue, M.J., 2012. Synthesizing phylogenetic knowledge for ecological research. *Ecology*, 93(sp8)
- Béguer-Pon, M., Castonguay, M., Shan, S., Benchetrit, J. and Dodson, J.J., 2015. Direct observations of American eels migrating across the continental shelf to the Sargasso Sea. *Nature Communications*, 6, p.8705.
- Bernard, A.M., Feldheim, K.A. and Shivji, M.S., 2012. Development and characterization of 11 novel microsatellite loci for the roundscale spearfish *Tetrapturus georgii* and their cross-species amplification among other istiophorid species. *Journal of fish biology*, 81(5), pp.1781-1786.
- Bernard, A.M., Feldheim, K.A., Richards, V.P., Nemeth, R.S. and Shivji, M.S., 2012. Development and characterization of fifteen novel microsatellite loci for the Nassau grouper (*Epinephelus striatus*) and their utility for cross-amplification on a suite of closely related species. *Conservation Genetics Resources*, 4(4), pp.983-986.
- Blair, J.E. and Hedges, S.B., 2005. Molecular phylogeny and divergence times of deuterostome animals. *Molecular Biology and Evolution*, 22(11), pp.2275-2284.,
- Bolnick, D.I. and Fitzpatrick, B.M., 2007. Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.*, 38, pp.459-487.
- Bracken, F.S., Hoelzel, A.R., Hume, J.B. and Lucas, M.C., 2015. Contrasting population genetic structure among freshwater-resident and anadromous lampreys: the role of demographic history, differential dispersal and anthropogenic barriers to movement. *Molecular Ecology*, 24(6), pp.1188-1204.
- Bracken, F.S., Hoelzel, A.R., Hume, J.B. and Lucas, M.C., 2015. Contrasting population genetic structure among freshwater-resident and anadromous lampreys: the role of

demographic history, differential dispersal and anthropogenic barriers to movement. *Molecular Ecology*, 24(6), pp.1188-1204.

Cairney, M., Taggart, J.B. and Høyheim, B., 2000. Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. *Molecular Ecology*, 9(12), pp.2175-2178.

Carlsson, J., McDowell, J.A.N., DÍAZ-JAIMES, P.Í.N.D.A.R.O., Carlsson, J.E., Boles, S.B., Gold, J.R. and Graves, J.E., 2004. Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranean Sea. *Molecular Ecology*, 13(11), pp.3345-3356.

Carrea, C., Paterson, I.G., Cussac, V.E. and Ruzzante, D.E., 2009. Ten novel microsatellite loci characterized for a remarkably widespread fish: *Galaxias maculatus* (Galaxiidae). *Molecular Ecology Resources*, 9(6), pp.1503-1505.

Carreras-Carbonell, J., Macpherson, E. and Pascual, M., 2006. Population structure within and between subspecies of the Mediterranean triplefin fish *Tripterygion delaisi* revealed by highly polymorphic microsatellite loci. *Molecular Ecology*, 15(12), pp.3527-3539.

Carvalho-Costa, L.F., Hatanaka, T. and Galetti, P.M., 2006. Isolation and characterization of polymorphic microsatellite markers in the migratory freshwater fish *Prochilodus costatus*. *Molecular Ecology Resources*, 6(3), pp.818-819.

Chakraborty, R., Kimmel, M., Stivers, D.N., Davison, L.J. and Deka, R., 1997. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proceedings of the National Academy of Sciences*, 94(3), pp.1041-1046.

Chiang, T.Y., Lee, T.W., Lin, F.J., Huang, K.H. and Lin, H.D., 2008. Isolation and characterization of microsatellite loci in the endangered freshwater fish *Varicorhinus alticorpus* (Cyprinidae). *Conservation Genetics*, 9(5), pp.1399-1401.

Chubb, A.L., Zink, R.M. and Fitzsimons, J.M., 1998. Patterns of mtDNA variation in Hawaiian freshwater fishes: the phylogeographic consequences of amphidromy. *Journal of Heredity*, 89(1), pp.8-16.

Closs, G.P., Hicks, A.S. and Jellyman, P.G., 2013. Life histories of closely related amphidromous and non-migratory fish species: a trade-off between egg size and fecundity. *Freshwater Biology*, 58(6), pp.1162-1177.

Coleman, R.A., Weeks, A.R. and Hoffmann, A.A., 2013. Balancing genetic uniqueness and genetic variation in determining conservation and translocation strategies: a

comprehensive case study of threatened dwarf galaxias, *Galaxiella pusilla* (Mack)(Pisces: Galaxiidae). *Molecular Ecology*, 22(7), pp.1820-1835.

Daemen, E., Cross, T., Ollevier, F. and Volckaert, F., 2001. Analysis of the genetic structure of European eel (*Anguilla anguilla*) using microsatellite DNA and mtDNA markers. *Marine Biology*, 139(4), pp.755-764.

D'Aloia, C.C., Bogdanowicz, S.M., Harrison, R.G. and Buston, P.M., 2014. Seascape continuity plays an important role in determining patterns of spatial genetic structure in a coral reef fish. *Molecular Ecology*, 23(12), pp.2902-2913.

Davies, C.A., Gosling, E.M., Was, A., Brophy, D. and Tysklind, N., 2011. Microsatellite analysis of albacore tuna (*Thunnus alalunga*): population genetic structure in the North-East Atlantic Ocean and Mediterranean Sea. *Marine biology*, 158(12), pp.2727-2740.

DiBlasi, E., Johnson, K.P., Stringham, S.A., Hansen, A.N., Beach, A.B., Clayton, D.H. and Bush, S.E., 2018. Phoretic dispersal influences parasite population genetic structure. *Molecular ecology*.

Dick, S., Shurin, J.B. and Taylor, E.B., 2014. Replicate divergence between and within sounds in a marine fish: the copper rockfish (*Sebastes caurinus*). *Molecular Ecology*, 23(3), pp.575-590.

Dudgeon, C.L., Feldheim, K., Schick, M. and Ovenden, J.R., 2006. Polymorphic microsatellite loci for the zebra shark *Stegostoma fasciatum*. *Molecular Ecology Resources*, 6(4), pp.1086-1088.

Douglas, M., Keck, B.P., Ruble, C., Petty, M., Shute, J.R., Rakes, P. and Hulsey, C.D., 2013. Pelagic larval duration predicts extinction risk in a freshwater fish clade. *Biology letters*, 9(6), p.20130672.

Ely, B., Viñas, J., Bremer, J.R.A., Black, D., Lucas, L., Covello, K., Labrie, A.V. and Thelen, E., 2005. Consequences of the historical demography on the global population structure of two highly migratory cosmopolitan marine fishes: the yellowfin tuna (*Thunnus albacares*) and the skipjack tuna (*Katsuwonus pelamis*). *BMC Evolutionary Biology*, 5(1), p.19.

Etnier, D.A. and Starnes, W.C., 1993. *The fishes of Tennessee*. University of Tennessee Press.

Feldheim, K.A., Gruber, S.H. and Ashley, M.V., 2001. Population genetic structure of the lemon shark (*Negaprion brevirostris*) in the western Atlantic: DNA microsatellite variation. *Molecular Ecology*, 10(2), pp.295-303.

- Feldheim, K.A., Kreiser, B.R., Schmidt, B., Duvernell, D.D. and Schaefer, J.F., 2014. Isolation and characterization of microsatellite loci for the blackstripe topminnow *Fundulus notatus* and their variability in two closely related species. *Journal of fish Biology*, 85(5), pp.1726-1732.
- Feldheim, K.A., Sanchez, P.J., Matamoros, W.A., Schaefer, J.F. and Kreiser, B.R., 2009. Isolation and characterization of microsatellite loci for mountain mullet (*Agonostomus monticola*). *Molecular Ecology Resources*, 9(6), pp.1482-1484.
- Feutry, P., Vergnes, A., Broderick, D., Lambourdière, J., Keith, P. and Ovenden, J.R., 2013. Stretched to the limit; can a short pelagic larval duration connect adult populations of an Indo-Pacific diadromous fish (*Kuhlia rupestris*)?. *Molecular Ecology*, 22(6), pp.1518-1530.
- Finger, A.J. and May, B., 2015. Conservation genetics of a desert fish species: the Lahontan tui chub (*Siphateles bicolor* ssp.). *Conservation Genetics*, 16(3), pp.743-758.
- Fisch, K.M., Henderson, J.M., Burton, R.S. and May, B., 2011. Population genetics and conservation implications for the endangered delta smelt in the San Francisco Bay-Delta. *Conservation Genetics*, 12(6), pp.1421-1434.
- Froese R, Pauly D. 2018. FishBase. Available from: www.fishbase.org, version March, 2018.
- Green, B.S. and Fisher, R., 2004. Temperature influences swimming speed, growth and larval duration in coral reef fish larvae. *Journal of Experimental Marine Biology and Ecology*, 299(1), pp.115-132.
- Guillemaud, T., Streiff, R., Santos, R.S., Afonso, P., Morato, T. and Cancela, M.L., 2000. Microsatellite characterization in the rainbow wrasse *Coris julis* (Pisces: Labridae). *Journal of Fish Biology*, 44, pp.213-232.
- Hardisty, M.W. and Potter, I.C. 1971. Paired species. In: Hardisty, M.W. and Potter, I.C., (eds.) *The biology of lampreys*. Academic Press, London, V1: 249-277.
- Harris, L.N. and Taylor, E.B., 2010. Pleistocene glaciations and contemporary genetic diversity in a Beringian fish, the broad whitefish, *Coregonus nasus* (Pallas): inferences from microsatellite DNA variation. *Journal of Evolutionary Biology*, 23(1), pp.72-86.
- Harris, L.N., Moore, J.S., Galpern, P., Tallman, R.F. and Taylor, E.B., 2014. Geographic influences on fine-scale, hierarchical population structure in northern Canadian populations of anadromous Arctic Char (*Salvelinus alpinus*). *Environmental biology of fishes*, 97(11), pp.1233-1252.

Harvey, M.G., Aleixo, A., Ribas, C.C. and Brumfield, R.T., 2017a. Habitat association predicts genetic diversity and population divergence in Amazonian birds. *The American Naturalist*, 190(5), pp.631-648.

Harvey, M.G., Seeholzer, G.F., Smith, B.T., Rabosky, D.L., Cuervo, A.M. and Brumfield, R.T., 2017b. Positive association between population genetic differentiation and speciation rates in New World birds. *Proceedings of the National Academy of Sciences*, 114(24), pp.6328-6333.

Hendry, A.P., Wenburg, J.K., Bentzen, P., Volk, E.C. and Quinn, T.P., 2000. Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science*, 290(5491), pp.516-518.

Ho, L. S. T. and Ane, C. (2014) A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Systematic Biology*, 63(3):397-408.

Hoarau, G., Rijnsdorp, A.D., Van der Veer, H.W., Stam, W.T. and Olsen, J.L., 2002. Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: microsatellites revealed large-scale spatial and temporal homogeneity. *Molecular Ecology*, 11(7), pp.1165-1176.

Hoareau, T.B., Bosc, P., Valade, P. and Berrebi, P., 2007. Gene flow and genetic structure of *Sicyopterus lagocephalus* in the south-western Indian Ocean, assessed by intron-length polymorphism. *Journal of Experimental Marine Biology and Ecology*, 349(2), pp.223-234.

Hogan, J.D., Blum, M.J. and Walter, R.P., 2011. Characterization of ten novel microsatellite markers in *Awaous guamensis* with comments on cross amplification in congeners and other amphidromous fish native to Hawai'i. *Conservation Genetics Resources*, 3(2), pp.275-277.

Huey, J.A., Espinoza, T. and Hughes, J.M., 2013. Regional panmixia in the mullet '*Mugil cephalus*' along the coast of Eastern Queensland; revealed using six highly polymorphic microsatellite loci. *Proceedings of the Royal Society of Queensland*, The, 118, p.7.

Hughes, J.M., Schmidt, D.J., Macdonald, J.I., Huey, J.A. and Crook, D.A., 2014. Low interbasin connectivity in a facultatively diadromous fish: evidence from genetics and otolith chemistry. *Molecular Ecology*, 23(5), pp.1000-1013.

Hutchinson, G.E. and MacArthur, R.H., 1959. A theoretical ecological model of size distributions among species of animals. *The American Naturalist*, 93(869), pp.117-125.

Jones, F.C., Chan, Y.F., Schmutz, J., Grimwood, J., Brady, S.D., Southwick, A.M., Absher, D.M., Myers, R.M., Reimchen, T.E., Deagle, B.E. and Schluter, D., 2012. A

genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current biology*, 22(1), pp.83-90.

Keefer, M.L. and Caudill, C.C., 2014. Homing and straying by anadromous salmonids: a review of mechanisms and rates. *Reviews in fish biology and fisheries*, 24(1), pp.333-368.

Kisel, Y., Moreno-Letelier, A.C., Bogarín, D., Powell, M.P., Chase, M.W. and Barraclough, T.G., 2012. Testing the link between population genetic differentiation and clade diversification in Costa Rican orchids. *Evolution: International Journal of Organic Evolution*, 66(10), pp.3035-3052.

Kochmer, J.P. and Wagner, R.H., 1988. Why are there so many kinds of passerine birds? Because they are small. A reply to Raikow. *Systematic Zoology*, 37(1), pp.68-69.

Koizumi, I., Yamamoto, S. and Maekawa, K., 2006. Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden charr. *Molecular Ecology*, 15(11), pp.3175-3189.

Kokita, T., Takahashi, S. and Kumada, H., 2013. Molecular signatures of lineage-specific adaptive evolution in a unique sea basin: the example of an anadromous goby *Leucopsarion petersii*. *Molecular ecology*, 22(5), pp.1341-1355.

Krieg, F., Estoup, A., Triantafyllidis, A. and Guyomard, R., 1999. Isolation of microsatellite loci in European catfish, *Silurus glanis*. *Molecular Ecology*, 8(11), pp.1964-1966.

Kumar, S., and S. B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature* 392:917–920

Leider, S.A., 1989. Increased straying by adult steelhead trout, *Salmo gairdneri*, following the 1980 eruption of Mount St. Helens. *Environmental Biology of Fishes*, 24(3), pp.219-229.

Luiz, O.J., Madin, J.S., Robertson, D.R., Rocha, L.A., Wirtz, P. and Floeter, S.R., 2012. Ecological traits influencing range expansion across large oceanic dispersal barriers: insights from tropical Atlantic reef fishes. *Proc. R. Soc. B*, 279(1730), pp.1033-1040.

Maggio, T., Lo Brutto, S., Garoia, F., Tinti, F. and Arculeo, M., 2009. Microsatellite analysis of red mullet *Mullus barbatus* (Perciformes, Mullidae) reveals the isolation of the Adriatic Basin in the Mediterranean Sea. *ICES Journal of Marine Science*, 66(9), pp.1883-1891.

McCraney, W.T., 2009. Rampant drift in the endangered tidewater goby (*Eucyclogobius newberryi*): comparing genetic variation of naturally and artificially fragmented populations (Doctoral dissertation, Humboldt State University).

McCulloch, G.A., Wallis, G.P. and Waters, J.M., 2009. Do insects lose flight before they lose their wings? Population genetic structure in subalpine stoneflies. *Molecular ecology*, 18(19), pp.4073-4087.

McDowall, R.M., 1970. *The galaxiid fishes of New Zealand*. Museum of Comparative Zoology, Harvard University.

McDowall, R.M., 1988. *Diadromy in fishes: migrations between freshwater and marine environments*. Croom Helm.

McDowall, R. M. 2001. Diadromy, diversity and divergence: implications for speciation processes in fishes. *Fish and Fisheries*, 2(3), 278-285.

McDowall, R.M., 2004. Ancestry and amphidromy in island freshwater fish faunas. *Fish and fisheries*, 5(1), pp.75-85.

McDowall, R.M., 2007. On amphidromy, a distinct form of diadromy in aquatic organisms. *Fish and fisheries*, 8(1), pp.1-13.

Melo, B.F., Sidlauskas, B.L., Hoekzema, K., Frable, B.W., Vari, R.P. and Oliveira, C., 2016. Molecular phylogenetics of the Neotropical fish family Prochilodontidae (Teleostei: Characiformes). *Molecular phylogenetics and evolution*, 102, pp.189-201.

Mesquita, N., Cunha, C., Hänfling, B., Carvalho, G.R., Ze-Ze, L., Tenreiro, R. and Coelho, M.M., 2003. Isolation and characterization of polymorphic microsatellite loci in the endangered Portuguese freshwater fish *Squalius aradensis* (Cyprinidae). *Molecular Ecology Resources*, 3(4), pp.572-574.

Miller, R.R., Minckley, W.L. and Norris, S.M., 2005. *Freshwater fishes of Mexico* (No. QL 629. M54 2005).

Moody, K.N., Hunter, S.N., Childress, M.J., Blob, R.W., Schoenfuss, H.L., Blum, M.J. and Ptacek, M.B., 2015. Local adaptation despite high gene flow in the waterfall-climbing Hawaiian goby, *Sicyopterus stimpsoni*. *Molecular Ecology*, 24(3), pp.545-563.

Moyer, G.R., Sloss, B.L., Kreiser, B.R. and Feldheim, K.A., 2009. Isolation and characterization of microsatellite loci for alligator gar (*Atractosteus spatula*) and their variability in two other species (*Lepisosteus oculatus* and *L. osseus*) of Lepisosteidae. *Molecular Ecology Resources*, 9(3), pp.963-966.

- Moyer, G.R., Sloss, B.L., Kreiser, B.R. and Feldheim, K.A., 2009. Isolation and characterization of microsatellite loci for alligator gar (*Atractosteus spatula*) and their variability in two other species (*Lepisosteus oculatus* and *L. osseus*) of Lepisosteidae. *Molecular Ecology Resources*, 9(3), pp.963-966.
- Moyer, G.R., Sloss, B.L., Kreiser, B.R. and Feldheim, K.A., 2009. Isolation and characterization of microsatellite loci for alligator gar (*Atractosteus spatula*) and their variability in two other species (*Lepisosteus oculatus* and *L. osseus*) of Lepisosteidae. *Molecular Ecology Resources*, 9(3), pp.963-966.
- Myers, G.S., 1949. Usage of anadromous, catadromous and allied terms for migratory fishes. *Copeia*, 1949(2), pp.89-97.
- Myers, R.A., Mertz, G. and Bridson, J., 1997. Spatial scales of interannual recruitment variations of marine, anadromous, and freshwater fish. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(6), pp.1400-1407.
- Nanninga, G.B., Saenz-Agudelo, P., Manica, A. and Berumen, M.L., 2014. Environmental gradients predict the genetic population structure of a coral reef fish in the Red Sea. *Molecular Ecology*, 23(3), pp.591-602.
- Neilson, M.E. and Stepien, C.A., 2011. Historic speciation and recent colonization of Eurasian monkey gobies (*Neogobius fluviatilis* and *N. pallasii*) revealed by DNA sequences, microsatellites, and morphology. *Diversity and Distributions*, 17(4), pp.688-702.
- Neilson, M.E. and Stepien, C.A., 2011. Historic speciation and recent colonization of Eurasian monkey gobies (*Neogobius fluviatilis* and *N. pallasii*) revealed by DNA sequences, microsatellites, and morphology. *Diversity and Distributions*, 17(4), pp.688-702.
- Olden, J.D., Hogan, Z.S. and Zanden, M.J.V., 2007. Small fish, big fish, red fish, blue fish: size-biased extinction risk of the world's freshwater and marine fishes. *Global Ecology and Biogeography*, 16(6), pp.694-701.
- O'reilly, P.T., Canino, M.F., Bailey, K.M. and Bentzen, P., 2004. Inverse relationship between FST and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. *Molecular Ecology*, 13(7), pp.1799-1814.
- Ohara, K., Hotta, M., Takahashi, D., Asahida, T., Ida, H. and Umino, T., 2009. Use of microsatellite DNA and otolith Sr: Ca ratios to infer genetic relationships and migration history of four morphotypes of *Rhinogobius* sp. OR. *Ichthyological research*, 56(4), pp.373-379.

Palkovacs, E.P., Hasselman, D.J., Argo, E.E., Gephard, S.R., Limburg, K.E., Post, D.M., Schultz, T.F. and Willis, T.V., 2014. Combining genetic and demographic information to prioritize conservation efforts for anadromous alewife and blueback herring. *Evolutionary Applications*, 7(2), pp.212-226.

Paradis, E., Claude, J. and Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20(2), pp.289-290.

Partridge, C., Boettcher, A. and Jones, A.G., 2012. Population structure of the gulf pipefish in and around Mobile Bay and the northern gulf of Mexico. *Journal of Heredity*, 103(6), pp.821-830.

Peters, R. H. 1983 The ecological implications of body size. Cambridge Studies in Ecology. Cambridge University Press.

Pirog, A., Jaquemet, S., Blaison, A., Soria, M. and Magalon, H., 2016. Isolation and characterization of eight microsatellite loci from *Galeocerdo cuvier* (tiger shark) and cross-amplification in *Carcharhinus leucas*, *Carcharhinus brevipinna*, *Carcharhinus plumbeus* and *Sphyrna lewini*. *PeerJ*, 4, p.e2041.

Palumbi, S.R. 1994. Genetic divergence, reproductive isolation, and marine speciation. *Annu. Rev. Ecol. Syst.* 25:547–572.

Palumbi, S.R. 2003. Population genetics, demographic connectivity, and the design of marine reserves. *Ecological applications*, 13(sp1), pp.146-158.

Quinn, T.P. and Fresh, K., 1984. Homing and straying in chinook salmon (*Oncorhynchus tshawytscha*) from Cowlitz River Hatchery, Washington. *Canadian Journal of Fisheries and Aquatic Sciences*, 41(7), pp.1078-1082.

Rabosky, D. L., Santini, F., Eastman, J., Smith, S. A., Sidlauskas, B., Chang, J. *et al.* Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat Commun.* 2013a;4: 1958.

Rabosky, D.L. and Matute, D.R., 2013. Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. *Proceedings of the National Academy of Sciences*, p.201305529.

Rousset, F., 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, 145(4), pp.1219-1228.

Rueda, E.C., Sommer, J., Scarabotti, P., Markariani, R. and Ortí, G., 2011. Isolation and characterization of polymorphic microsatellite loci in the migratory freshwater fish Ryman, N., Allendorf, F.W. and Ståhl, G., 1979. Reproductive isolation with little

genetic divergence in sympatric populations of brown trout (*Salmo trutta*). *Genetics*, 92(1), pp.247-262.

Prochilodus lineatus (Characiformes: Prochilodontidae). *Conservation Genetics Resources*, 3(4), pp.681-684.

Sanches, A. and GALETTI, P., 2006. Microsatellites loci isolated in the freshwater fish *Brycon hilarii*. *Molecular Ecology Resources*, 6(4), pp.1045-1046.

Schmidt, D.J., Crook, D.A., Macdonald, J.I., Huey, J.A., Zampatti, B.P., Chilcott, S., Raadik, T.A. and Hughes, J.M., 2013. Migration history and stock structure of two putatively diadromous teleost fishes, as determined by genetic and otolith chemistry analyses. *Freshwater Science*, 33(1), pp.193-206.

Schmidt, D.J., Crook, D.A., O'connor, J.P. and Hughes, J.M., 2011. Genetic analysis of threatened Australian grayling *Prototroctes maraena* suggests recruitment to coastal rivers from an unstructured marine larval source population. *Journal of Fish Biology*, 78(1), pp.98-111.

Schug, M.D., Hutter, C.M., Wetterstrand, K.A., Gaudette, M.S., Mackay, T.F. and Aquadro, C.F., 1998. The mutation rates of di-, tri- and tetranucleotide repeats in *Drosophila melanogaster*. *Molecular Biology and Evolution*, 15(12), pp.1751-1760.

Schultz, J.K., Feldheim, K.A., Gruber, S.H., Ashley, M.V., McGovern, T.M. and Bowen, B.W., 2008. Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). *Molecular Ecology*, 17(24), pp.5336-5348.

Selkoe, K.A., Watson, J.R., White, C., Horin, T.B., Iacchei, M., Mitarai, S., Siegel, D.A., Gaines, S.D. and Toonen, R.J., 2010. Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology*, 19(17), pp.3708-3726.

Sekino, M. and Hara, M., 2000. Isolation and characterization of microsatellite DNA loci in Japanese flounder *Paralichthys olivaceus* (Pleuronectiformes, Pleuronectoidei, Paralichthyidae). *Molecular Ecology*, 9(12), pp.2201-2203.

Sekino, M., Takagi, N., Hara, M. and Takahashi, H., 2000. Microsatellites in rockfish *Sebastes thompsoni* (Scorpaenidae). *Molecular Ecology*, 9(5), pp.634-636.

Shaddick, K., Gilligan, D.M., BurrIDGE, C.P., Jerry, D.R., Truong, K. and Beheregaray, L.B., 2011. Historic divergence with contemporary connectivity in a catadromous fish, the estuary perch (*Macquaria colonorum*). *Canadian Journal of Fisheries and Aquatic Sciences*, 68(2), pp.304-318.

- Shaw, P.W., Turan, C., Wright, J.M., O'connell, M. and Carvalho, G.R., 1999. Microsatellite DNA analysis of population structure in Atlantic herring (*Clupea harengus*), with direct comparison to allozyme and mtDNA RFLP analyses. *Heredity*, 83(4), p.490.
- Smith, G.R., 1992. Phylogeny and biogeography of the Catostomidae, freshwater fishes of North America and Asia. *Systematics, historical ecology, and North American freshwater fishes*, pp.778-826.
- Strathmann, R.R., 1985. Feeding and nonfeeding larval development and life-history evolution in marine invertebrates. *Annual review of ecology and systematics*, 16(1), pp.339-361.
- Strona, G., Galli, P., Montano, S., Seveso, D. and Fattorini, S., 2012. Global-scale relationships between colonization ability and range size in marine and freshwater fish. *PLoS One*, 7(11), p.e49465.
- Switzer, J.F., Welsh, S.A. and King, T.L., 2008. Microsatellite DNA primers for the candy darter, *Etheostoma osburni* and variegate darter, *Etheostoma variatum*, and cross-species amplification in other darters (Percidae). *Molecular Ecology Resources*, 8(2), pp.335-338.
- Switzer, J.F., Welsh, S.A. and King, T.L., 2008. Microsatellite DNA primers for the candy darter, *Etheostoma osburni* and variegate darter, *Etheostoma variatum*, and cross-species amplification in other darters (Percidae). *Molecular Ecology Resources*, 8(2), pp.335-338.
- Takagi, M., Shoji, E. and Taniguchi, N., 1999. Microsatellite DNA polymorphism to reveal genetic divergence in ayu, *Plecoglossus altivelis*. *Fisheries science*, 65(4), pp.507-512.
- Tang, S.J., Liu, Z.Z., Tang, W.Q. and Yang, J.Q., 2009. A simple method for isolation of microsatellites from the mudskipper (*Boleophthalmus pectinirostris*), without constructing a genomic library. *Conservation genetics*, 10(6), p.1957.
- Thériault, V., Garant, D., Bernatchez, L. and Dodson, J.J., 2007. Heritability of life-history tactics and genetic correlation with body size in a natural population of Brook Charr (*Salvelinus fontinalis*). *Journal of evolutionary biology*, 20(6), pp.2266-2277.
- Tripp-Valdez, M.A., de León, F.J.G., Ortega-García, S., Lluch-Cota, D., López-Martínez, J. and Cruz, P., 2010. Population genetic structure of dolphinfish (*Coryphaena hippurus*) in the Gulf of California, using microsatellite loci. *Fisheries Research*, 105(3), pp.172-177.

- Tseng, M.C., Tzeng, W.N. and Lee, S.C., 2006. Population genetic structure of the Japanese eel *Anguilla japonica* in the northwest Pacific Ocean: evidence of non-panmictic populations. *Marine Ecology Progress Series*, 308, pp.221-230.
- Vera-Escalona, I., Paterson, I.G., Habit, E. and Ruzzante, D.E., 2014. Development and characterization of 15 novel microsatellite markers for the freshwater fish *Galaxias platei*. *Conservation genetics resources*, 6(4), pp.899-901.
- Waples, R. S. 1987. A multispecies approach to the analysis of gene flow in marine shore fishes. *Evolution* 41:385–400
- Waters, J.M., Esa, Y.B. and Wallis, G.P., 1999. Characterization of microsatellite loci from a New Zealand freshwater fish (*Galaxias vulgaris*) and their potential for analysis of hybridization in *Galaxiidae*. *Molecular Ecology*, 8(6), pp.1080-1082.
- Weir, B.S. and Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *evolution*, 38(6), pp.1358-1370.
- Wenburg, J.K., Bentzen, P. and Foote, C.J., 1998. Microsatellite analysis of genetic population structure in an endangered salmonid: the coastal cutthroat trout (*Oncorhynchus clarki clarki*). *Molecular Ecology*, 7(6), pp.733-749.
- Wilson, A.B., 2006. Genetic signature of recent glaciation on populations of a near-shore marine fish species (*Syngnathus leptorhynchus*). *Molecular Ecology*, 15(7), pp.1857-1871.
- Wirth, T., Saint-Laurent, R. and Bernatchez, L., 1999. Isolation and characterization of microsatellite loci in the walleye (*Stizostedion vitreum*), and cross-species amplification within the family Percidae. *Molecular Ecology*, 8(11), pp.1960-1962.
- Wright, S., 1943. Isolation by distance. *Genetics*, 28(2), p.114.
- Yang, J.Q., Zhou, X.D., Liu, D., Liu, Z.Z. and Tang, W.Q., 2011. Isolation and characterization of microsatellite loci in the fish *Coilia mystus* (Clupeiformes: Engraulidae) using PCR-based isolation of microsatellite arrays. *Genet Mol Res*, 10(3), pp.1514-1517.
- Yue, G.H., Zhu, Z.Y., Lo, L.C., Wang, C.M., Lin, G., Feng, F., Pang, H.Y., Li, J., Gong, P., Liu, H.M. and Tan, J., 2009. Genetic variation and population structure of Asian seabass (*Lateolabrax japonicus*) in the Asia-Pacific region. *Aquaculture*, 293(1-2), pp.22-28.

Appendix A – Figures and Tables

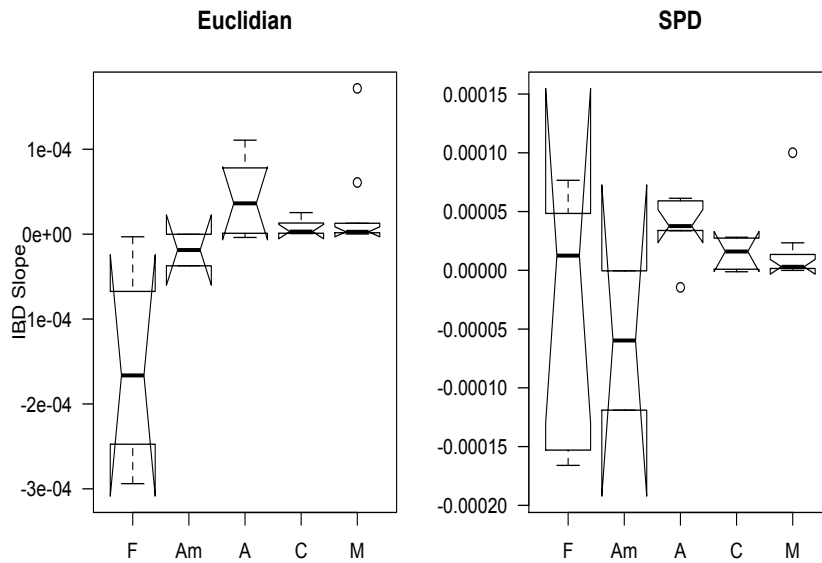


Figure 2.1 IBS slopes arranged by life history
Distribution of slopes of genetic variation vs. Euclidian and. shortest path distance

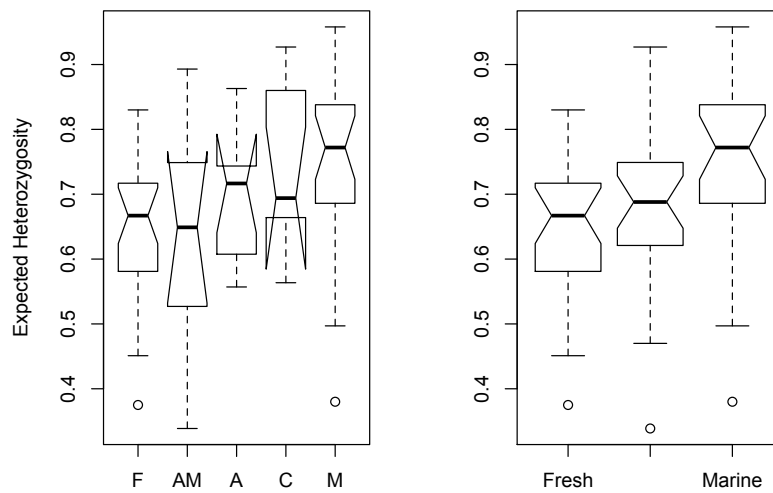


Figure 2.2 Distribution of Heterozygosity.
Heterozygosity (H_e) of freshwater (F), amphidromous (AM), Anadromous (A), Catadromous(C), Pooled Diadromous (D = AM + A + C), and Marine (M)

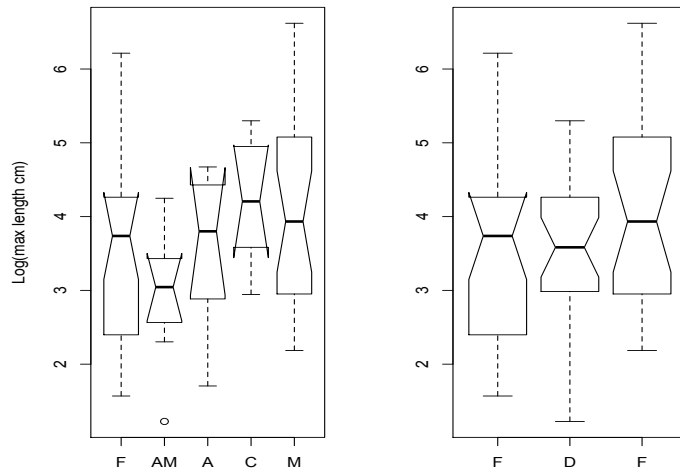


Figure 2.3 Distribution of maximum length.
Maximum Length (CM) of freshwater (F), amphidromous (AM), anadromous (A),
catadromous(C), pooled diadromous (D = AM + A + C), and Marine (M)

Table 2.1 Heterozygosity

Species	LH	H _e	max length (CM)	Source
<i>Siphateles bicolor</i>	F	0.758	45	Finger and May 2015
<i>Lampetra planeri</i>	F	0.532	20.5	Bracken <i>et al.</i> 2015
<i>Hypomesus transpacificus</i>	F	0.83	12	Fisch <i>et al.</i> 2011
<i>Squalius aradensis</i>	F	0.551	13.1	Mesquita <i>et al.</i> 2003
<i>Brycon hilarii</i>	F	0.667	56	Sanches and Galetti 2006
<i>Prochilodus costatus</i>	F	0.64	42	Carvalho-Costa <i>et al.</i> 2006
<i>Varicorhinus alticorpus</i>	F	0.821	50	Chiang, <i>et al.</i> 2008
<i>Prochilodus lineatus</i>	F	0.649	80	Rueda <i>et al.</i> 2011
<i>Fundulus notatus</i>	F	0.712	8	Feldheim <i>et al.</i> 2014
<i>Fundulus olivaceus</i>	F	0.748	8	Feldheim <i>et al.</i> 2014
<i>Fundulus euryzonus</i>	F	0.712	9	Feldheim <i>et al.</i> 2014
<i>Atractosteus spatula</i>	F	0.5	305	Moyer <i>et al.</i> 2009
<i>Lepisosteus oculatus</i>	F	0.654	50	Moyer <i>et al.</i> 2009
<i>Lepisosteus osseus</i>	F	0.637	200	Moyer <i>et al.</i> 2009
<i>Stizostedion vitreum</i>	F	0.705	107	Wirth <i>et al.</i> 1999
<i>Silurus glanis</i>	F	0.608	500	Krieg <i>et al.</i> 1999
<i>Silurus triostegus</i>	F	0.766	99	Krieg <i>et al.</i> 1999
<i>Silurus aristotelis</i>	F	0.717	46	Krieg <i>et al.</i> 1999
<i>Neogobius fluviatilis</i>	F	0.451	20	Neilson and Stepien 2011
<i>Neogobius pallasii</i>	F	0.77	20	Neilson and Stepien 2011
<i>Etheostoma osburni</i>	F	0.581	10	Switzer <i>et al.</i> 2008
<i>Etheostoma variatum</i>	F	0.706	11	Switzer <i>et al.</i> 2008
<i>Galaxias vulgaris</i>	F	0.685	10	Waters <i>et al.</i> 1999
<i>Galaxiella pusilla</i>	F	0.375	4.8	Coleman <i>et al.</i> 2013
<i>Fundulus heteroclitus</i>	M	0.497	15	Adams <i>et al.</i> 2006
<i>Thunnus alalunga</i>	M	0.799	140	Davies <i>et al.</i> 2011
<i>Thunnus thynnus</i>	M	0.681	458	Carlsson <i>et al.</i> 2004
<i>Boleophthalmus pectinirostris</i>	M	0.759	20	Tang <i>et al.</i> 2009
<i>Sebastes thompsoni</i>	M	0.691	30	Sekino <i>et al.</i> 2000
<i>Sebastes caurinus</i>	M	0.614	58	Dick <i>et al.</i> 2014
<i>Negaprion brevirostris</i>	M	0.785	340	Feldheim <i>et al.</i> 2001
<i>Mullus barbatus</i>	M	0.835	33.2	Maggio <i>et al.</i> 2009
<i>Eucyclogobius newberryi</i>	M	0.313	5.7	McCraney 2009
<i>Amphiprion bicinctus</i>	M	0.754	14	Nanninga <i>et al.</i> 2014
<i>Theragra chalcogramma</i>	M	0.85	91	O'Reilly <i>et al.</i> 2004
<i>Syngnathus leptorhynchus</i>	M	0.584	38.5	Wilson 2006
<i>Clupea harengus</i>	M	0.916	45	Shaw <i>et al.</i> 1999
<i>Tetrapturus georgii</i>	M	0.677	184	Bernard <i>et al.</i> 2012
<i>Epinephelus striatus</i>	M	0.804	122	Bernard <i>et al.</i> 2012
<i>Galeocerdo cuvier</i>	M	0.38	750	Pirog <i>et al.</i> 2016.
<i>Stegostoma fasciatum</i>	M	0.742	354	Dudgeon <i>et al.</i> 2006
<i>Coris julis</i>	M	0.57	30	Guillemaud <i>et al.</i> 2000
<i>Syngnathus scovelli</i>	M	0.917	18.3	Partridge <i>et al.</i> 2012
<i>Tripterygion delaisi</i>	M	0.707	8.9	Carreras-Carbonell <i>et al.</i> 2006
<i>Tripterygion xanthosoma</i>	M	0.841	8.9	Carreras- <i>et al.</i> 2006
<i>Coryphaena hippurus</i>	M	0.83	210	Tripp-Valdez <i>et al.</i> 2010

Table 2.1 Continued.

Species	LH	H _e	max length	
			(CM)	Source
<i>Symphodus ocellatus</i>	M	0.899	12	Arigoni and Largiader 2000
<i>Paralichthys olivaceus</i>	M	0.8	103	Sekino and Hara 2000
<i>Coregonus nasus</i>	A	0.569	71	Harris and Taylor 2010
<i>Lampetra fluviatilis</i>	A	0.621	50	Bracken <i>et al.</i> 2015
<i>Salvelinus alpinus</i>	A	0.749	107	Harris <i>et al.</i> 2014
<i>Alosa pseudoharengus</i>	A	0.594	40	Palkovacs <i>et al.</i> 2014
<i>Alosa aestivalis</i>	A	0.557	40	Palkovacs <i>et al.</i> 2014
<i>Lovettia sealii</i>	A	0.738	8	Schmidt <i>et al.</i> 2013
<i>Oncorhynchus clarkii clarkii</i>	A	0.71	99	Wenburg <i>et al.</i> 1998
<i>Leucopsarion petersii</i>	A	0.863	5.5	Kokita <i>et al.</i> 2013
<i>Salmo salar</i>	A	0.723	71	Cairney <i>et al.</i> 2000
<i>Mugil Cephalus</i>	C	0.927	100	Huey <i>et al.</i> 2013
<i>Anguilla anguilla</i>	C	0.682	133	Daemen <i>et al.</i> 2001
<i>Pseudaphritis urvillii</i>	C	0.855	36	Schmidt <i>et al.</i> 2013
<i>Macquaria colonorum</i>	C	0.022	75	Shaddick <i>et al.</i> 2011
<i>Anguilla japonica</i>	C	0.865	150	Tseng <i>et al.</i> 2006
<i>Lates calcarifer</i>	C	0.7	200	Yue <i>et al.</i> 2009
<i>Agonostomus monticola</i>	C	0.646	36	Feldheim <i>et al.</i> 2009
<i>Kuhlia rupestris</i>	C	0.688	45	Feutry <i>et al.</i> 2013
<i>Galaxias maculatus</i>	C	0.564	19	Carrea <i>et al.</i> 2009
<i>Sicyopterus stimpsoni</i>	AM	0.893	19.8	Moody <i>et al.</i> 2015
<i>Rhinogobius rubromaculatus</i>	AM	0.64	3.4	Ohara <i>et al.</i> 2009
<i>Plecoglossus altivelis altivelis</i>	AM	0.784	70	Takagi <i>et al.</i> 1999
<i>Awaous guamensis</i>	AM	0.47	24.5	Hogan <i>et al.</i> 2011
<i>Prototroctes maraena</i>	AM	0.652	33	Schmidt <i>et al.</i> 2011
<i>Sicyopterus lagocephalus</i>	AM	0.527	13	Hoareau <i>et al.</i> 2007
<i>Galaxias platei</i>	AM	0.339	30.9	Vera-Escalona <i>et al.</i> 2014
<i>Retropinna semoni</i>	AM	0.649	10	Hughes <i>et al.</i> 2014
<i>Coilia mystus</i>	AM	0.749	21	Yang <i>et al.</i> 2011

Reported expected heterozygosity (H_e) for each species. Life histories (LH) include marine (M), anadromous (A), catadromous (C), amphidromous (AM), and freshwater (F)

Table 2.2 Isolation by distance linear models.

Model	Pagel's λ	AIC	Δ AIC
<u>Euclidean</u>			
slope~ 1 (Null model)	1.56E-08	-283.5	0
slope~ Log(max length)	2.69E-09	-281.5507	1.9492
slope~ LH	2.81E-09	-278.7458	4.7542
slope~ LH + log(max length)	2.79E-09	-276.8207	6.6793
slope~ log(max length) * LH	2.40E-09	-269.0819	14.4181
<u>SPD</u>			
slope~ log(max length)	2.69E-09	-363.2265	0
slope~ 1 (Null model)	2.15E-08	-363.0	0.2265
slope~ LH	2.81E-09	-357.4028	5.8236
slope~ LH + log(max length)	2.79E-09	-358.9501	4.2763
slope~ log(max length) * LH	2.40E-09	-356.6887	6.5377

Models using life history and max length to explain the slop of Euclidian or SPD regressed against genetic variation between each population with the inferred phylogeny.

Table 2.3 Isolation by distance tables

Species	SPD	r	p-value	Species	SPD	r	p-value
Marine				Anadromous			
<i>S. leptorhynchus</i>	1.35E-05	0.9388	0.333	<i>L. fluviatilis</i>	-1.45E-05	-0.2754	0.94
<i>T. delaisi</i>	2.34E-05	0.3245	0.143	<i>L. sealii</i>	6.05E-05	0.719	0.033
<i>T. chalcogramma</i>	1.74E-07	0.2687	0.21	<i>C. nasus</i>	6.13E-05	0.6907	0.004
<i>A. bicinctus</i>	1.18E-05	0.7291	0.001	<i>L. petersii</i>	3.42E-05	0.9477	0.045
<i>E. newberryi</i>	3.56E-03	0.5174	0.006	<i>A. oxyrinchus</i>	3.36E-05	0.0432	0.375
<i>C. harengus</i>	1.56E-06	0.5754	0.042	<i>S. alpinus</i>	6.06E-05	0.2814	0.044
<i>F. heteroclitus</i>	1.00E-04	0.2895	0.032	<i>A. pseudoharengus</i>	5.76E-05	0.5903	0.001
<i>M. barbatus</i>	-8.03E-08	-0.0744	0.648	<i>A. aestivalis</i>	3.95E-05	0.5203	0.167
<i>T. alalunga</i>	3.00E-06	0.3241	0.113	Amphidromous			
<i>N. brevirostis</i>	2.12E-06	0.7676	0.166	<i>S. stimpsoni</i>	4.51E-07	0.4799	0.25
				<i>Rhinogobius sp.</i>	-1.19E-04	-0.391	0.969
Catadromous				Freshwater			
<i>L. calcarifer</i>	2.82E-05	0.8588	0.125	<i>L. planeri</i>	1.25E-05	0.0748	0.511
<i>K. rupestri</i>	1.14E-05	0.8422	0.026	<i>H. transpacificus</i>	-1.53E-04	-0.7126	1
<i>A. japonicus</i>	8.97E-07	0.0977	0.362	<i>N. pallasii</i>	-1.66E-04	-0.3210	0.957
<i>M. colonorum</i>	2.74E-05	0.2943	0.035	<i>N. fluviatilis</i>	4.84E-05	0.0392	0.374
<i>P. urvillii</i>	2.06E-05	0.5363	0.048	<i>S. bicolor ssp.</i>	7.66E-05	0.4219	0.183
<i>A. anguilla</i>	-1.22E-06	-0.4053	1				

Slopes of linear regression by species estimated from genetic differentiation ($1/1-F_{ST}$) distance and Shortest path distance (SPD), correlation coefficient (r), and associated p-value.

Table 2.4 Heterozygosity models

Model	Pagel's λ	AIC	Δ AIC
He~ 1 (Null model)	2.52E-09	-61.33	0
He~ log(max length)	2.69E-09	-59.96	1.37
He~ LH	2.81E-09	-56.45	4.88
He~ LH + log(max length)	2.79E-09	-56.16	5.17
He~ log(max length) * LH	2.40E-09	-50.11	11.22

Model selection for Phylogenetic linear regression assessing Heterozygosity (H_e) as a function of life history (LH) and log(maximum length).

Appendix B – Additional Files

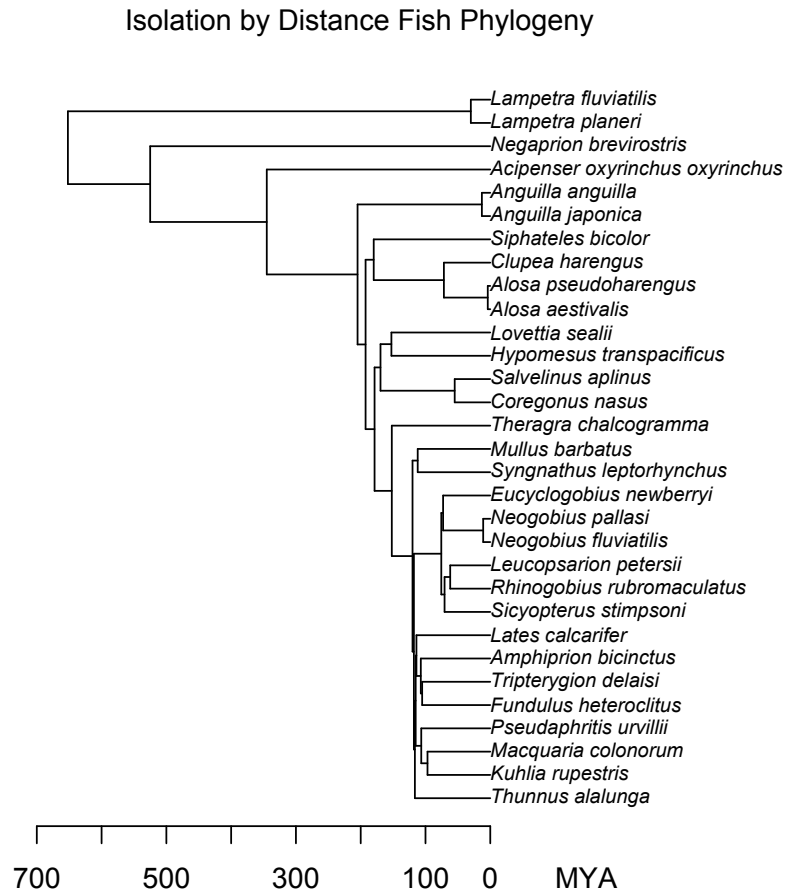


Figure A2.1
Reconstructed phylogeny used for IBD analyses

Heterozygosity Fish Phylogeny

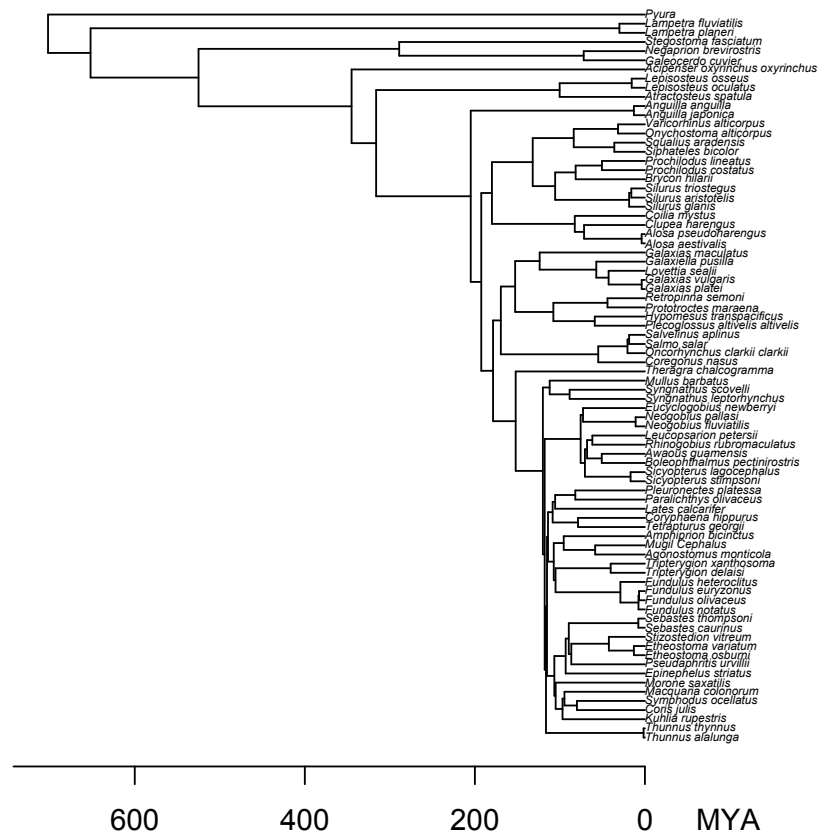


Figure A2.2
 Reconstructed phylogeny used for gene diversity analyses

Table A2.1
Populations locations

Species / Study / LH	Latitude	Longitude	Population ID
<i>L. fluviatilis</i>	53.980282	-1.318319	1LF
Bracken et al. 2015	54.357883	-1.549775	3LF
Anadromous	54.097477	-1.395795	5LF
	53.991965	-0.91404	7LF
	53.144795	-0.791247	9LF
	54.780365	-1.576352	10LF
	53.18648	-2.887331	12LF
	56.054808	-4.454507	14LF
	56.054809	-4.454517	15LF
	54.755192	-6.464189	17LF
	51.007225	3.752224	18LF
<i>L. planeri</i>	54.077314	-1.746886	2LP
Bracken et al. 2015	54.421152	-1.687299	4LP
Freshwater	54.216991	-1.724154	6LP
	54.237026	-1.042477	8LP
	54.722024	-1.943673	11LP
	52.926472	-3.082937	13LP
	56.054814	-4.454491	16LP
<i>Sicyopterus stimpsoni</i>	19.756267	-155.092038	Maili-09
Moody et al 2015	19.900105	-155.129298	Hakalau-09
Amphidromous	22.044453	-159.335937	Wailua-09
	21.951354	-159.666128	Waimea-09
	19.756567	-155.09204	Maili-10
	19.900184	-155.128872	Hakalau-10
	22.044296	-159.335414	Wailua-10
	21.951016	-159.666119	Waimea-10
	22.209124	-159.597778	Hanakapi'ai-10
	19.756567	-155.09204	Maili-11
	19.900184	-155.128872	Hakalau-11
	19.928208	-155.155579	Nanue-11
	22.044296	-159.335414	Wailua-11
	21.951016	-159.666119	Waimea-11
	22.209124	-159.597778	Hanakapi'ai-11
<i>Lates calcarifer</i>	12.346251	101.200468	Tha
Yue Et al 2009	8.845358	101.208219	Mas
Catadromous	2.61511	105.72555	Sing 1
	0.055692	105.692886	Ind
<i>Alosa pseudoharengus</i>	44.646134	-67.340379	1AP
Palkovacs et al. 2013	44.003067	-69.237515	2AP
Catadromous	44.469109	-68.441926	3AP
	43.855468	-69.566952	4AP
	43.978067	-69.854097	5AP

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
	43.063741	-70.70837	6AP
	42.982315	-70.94589	7AP
	43.065901	-70.91402	8AP
	43.008236	-70.856213	9AP
	42.752657	-70.814714	10AP
	42.350476	-71.024205	11AP
	41.52619	-72.061159	21AP
	41.300516	-72.23724	22AP
	41.309638	-72.349077	24AP
	41.835244	-72.809185	25AP
	41.299927	-72.899788	26AP
	41.392782	-73.075737	27AP
	41.22423	-73.111652	28AP
	41.027928	-73.594678	29AP
	40.712187	-74.02721	30AP
	39.6812808	-75.8062051	31AP
	38.5493326	-75.6978713	32AP
	38.112871	-77.0507	33AP
	37.265029	-76.5570336	34AP
	37.120939	-76.642279	35AP
	36.035937	-76.683706	36AP
	35.939109	-76.710658	37AP
	35.668878	-76.034628	38AP
<i>Alosa aestivalis</i>	44.646134	-67.340379	1AA
Palkovacs et al. 2013	44.003067	-69.237515	2AA
Catadromous	42.350476	-71.024205	7AA
	41.52619	-72.061159	10AA
	41.309638	-72.349077	11AA
	41.835244	-72.809185	12AA
	41.392782	-73.075737	13AA
	41.027928	-73.594678	14AA
	40.712187	-74.02721	15AA
	39.6812808	-75.8062051	16AA
	38.5493326	-75.6978713	17AA
	38.112871	-77.0507	18AA
	37.120939	-76.642279	19AA
	36.035971	-76.683555	20AA
	35.939092	-76.710722	21AA
	35.109363	-77.028934	22AA
	33.892387	-78.000835	23AA
	33.125831	-79.247859	24AA
	32.777745	-79.904608	25AA

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
	32.036466	-80.881775	26AA
	31.319229	-81.299833	27AA
	30.40112	-81.403639	28AA
<i>Syngnathus leptorhynchus</i>	57.029802	-135.448188	AK
Wilson 2006	48.550955	-122.538487	WA
Marine	44.620166	-124.07121	OR
	32.776334	-117.264809	CA
<i>Kuhlia rupestris</i>	-12.830724	45.207124	Mayotte
Feutry et al. 2013	-16.790942	49.827644	Madagascar
Catadromous	-21.04589	55.785985	Reunion
	26.070812	127.635641	Okinawa
	14.46277	122.297241	Philippines
	-15.800248	167.190921	vanuatu
	-21.639597	165.423031	New-Caledonia
	-16.291677	145.452653	Queensland_North
	-22.493981	150.74871	Queensland_Central
	-27.141931	153.090215	Queensland_South
<i>Tripterygion delaisi</i>	42.338182	3.247855	Cap_de_Creus
<i>xanthosoma</i>			
Carreras-Carbonell et al 2006	41.71983	2.936687	Tossa
	41.670091	2.790258	Blanes
	39.895595	0.68178	Columbretes
	38.696709	1.385574	Formentera
	37.633966	-0.689304	Cabo_de_Palos
	36.721471	-2.188296	Cabo_de_Gata
	36.006488	-5.60689	Tarifa
<i>Theragra chalcogramma</i>	42.25	142.5	Jpn98
O'Reilly et al. 2004	61.813	-178.5	NCBS97
Marine	54.413	-165.728	Uni97A
	54.45	-162.278	Uni97B
	54.331	-165.385	Uni98
	57.988	-154.212	Shel97
	57.598	-154.233	Shel98
	60.083	-148.317	PWS97
	60.083	-148.333	PWS98
	48.12	-122.77	PS98
<i>Anguilla japonica</i>	25.17653	121.413299	Tanshui_Taiwan
Tseng et al 2006	22.375039	120.57776	Fangliao_Taiwan
Catadromous	39.869137	124.22015	Yalu_River_China
	30.469639	121.567115	Hangzho_China
	23.320834	116.769211	Shantou_China

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
	36.350831	126.55578	Daecheon_myon_Korea
	34.769941	137.254856	Mikawa_Bay_Japan
<i>Macquaria colonorum</i>	-28.87507	153.58865	Richmond
shaddick <i>et al.</i> 2011	-29.425755	153.356807	Clarence
Catadromous	-33.563834	151.284863	Hawkesbury
	-34.903947	150.758193	Shoalhaven
	-35.711418	150.196417	Clyde
	-37.56224	149.771951	Mallacoota
	-37.781173	149.020864	Bemm
	-37.801419	148.545333	Snowy
			Gippslands
	-38.382234	147.185733	Merrimans
	-38.673937	146.646244	Albert
	-38.643039	145.73067	Tarwin
	-38.494852	145.429751	Bass
	-38.763611	143.674965	Barham
	-38.404045	142.508644	Hopkins
	-38.060215	140.988342	Glenelg
	-41.055855	144.657136	Arthur
<i>Lovettia sealii</i>	-41.166448	146.082342	Leven
Schmidt <i>et al.</i> 2007	-41.152299	146.550919	Rubicon
Anadromous	-43.260478	147.097289	Huon
	-42.916037	147.383142	Derwent
	-40.836685	145.316379	Black
	-38.686002	145.830881	Tarwin
<i>Pseudaphritis urvillii</i>	-35.526063	138.808826	Goolwa
Schmidt <i>et al.</i> 2007	-38.219777	141.77114	Darlots
Catadromous	-38.665085	145.944119	Tarwin
	-38.140006	147.078768	Thomson
	-37.71041	148.451729	Snowy
	-41.046126	147.625182	Great Forester
	-37.607581	148.901564	Bemm
	-38.384907	142.587709	Hopkins
<i>Coregonus nasus</i>	67.249999	-134.883334	PeelR
Harris and taylor 2010	66.978321	-133.25953	Arctic_RedR
Anadromous	67.61266	-134.121034	Point_Separation
	66.648156	-129.41791	Fort_Good_Hope
	68.213055	-133.466356	Campbell_Lake
	67.599999	-131.849999	Travaillant_River_South
	67.750001	-131.84998	Travaillant_River_North
	65.146937	-152.679278	Yukon_River
	65.502024	-150.189417	Yukon_River_Rampart_Rapids

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
	64.973301	-150.566723	Tanana_River
	70.819636	-154.14684	Teshekpuk_Lake
	66.601112	-160.329263	Selawik_River
	60.942684	-154.984659	Whitefish_Lake
	69.086021	55.128488	Pechora_River
<i>Amphiprion bicinctus</i>	27.909838	35.065333	Burcan
Nanninga et al.2014	27.138722	35.754192	An&Numan
Marine	26.624855	36.095918	Nuwayshiziyah
	25.582224	36.548775	Mashabi
	24.722586	37.150614	Abu&Matari
	24.149456	37.67505	Yanbu
	22.742487	38.782617	Shi'b&al&Bayda
	22.274669	39.048464	Haitham&Hai
	22.06962	38.771679	Abu&Madafi
	22.223754	38.968799	Palace&Reef
	21.814643	38.83754	Eagle&Reef
	21.676452	38.841322	Abu&Terr
	20.647445	39.394853	Tawil&Ral
	20.368743	39.650592	Um&Haj
	19.890366	39.960666	CanyonReef
	19.108864	40.489151	AQ3
	18.272823	40.736447	Maghabiyah
	17.78734	41.441833	Sumayr
	16.618166	41.93785	Farasan
<i>Sebastes caurinus</i>	49.190554	-124.820313	BSHI
Dick et al 2014	48.793023	-125.216668	BSC
Marine	49.225553	-125.59861	CSHI
	49.12472	-125.968054	CSC
	49.634118	-126.075083	NSHI
	49.579153	-126.670844	NSC
	50.183326	-127.3014	KSHI
	49.988885	-127.419722	KSC
	50.633855	-127.948333	QSHI
	50.45111	-128.062501	QSC
<i>Eucyclogobius newberryi</i>	41.831976	-124.232528	Lake_Earl
McCraney, et al 2010	41.247222	-124.105111	Stone_Lagoon
Marine	41.181443	-124.129083	Big_Lagoon
	39.471804	-123.805277	Virgin_Creek
	39.459524	-123.81011	Pudding
	40.859182	-124.102083	McDaniel
	40.849682	-124.081429	Gannon
	40.847878	-124.081229	Gannon

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
	40.843472	-124.082715	Jacoby
	40.785925	-124.10378	Wood_creek
	40.755718	-124.191034	Elk_River
	40.684442	-124.223569	Salmon_Creek
	40.647821	-124.307994	Eel_River
<i>Clupea harengus</i>	64.549999	-12.316667	IC_IC1
Shaw et al 2999	71.083332	29.166666	NS1_NW!
Marine	68.749999	-9.833333	NS2_NW2
	69.499999	-19.666666	BF_NW3
	49.583332	-124.666667	PC_PC
<i>Fundulus heteroclitus</i>	44.374502	-64.512324	Bridgewater_NS
Adams et al. 2006	43.928726	-69.711438	Chewonki_ME
Marine	43.319566	-70.556889	Wells_ME
	41.720916	-70.335931	Barnstable_MA
	41.566795	-70.915494	NewBedford_MA
	41.361111	-71.484692	Point_Judith_RI
	41.268226	-72.522538	Clinton_CN
	40.688074	-74.115229	Newark_Bay_NJ
	39.534284	-74.3514	Tuckerton_NJ
	39.03492	-74.781745	Stone_Harbor_NJ
	37.176624	-75.94113	Magotha_VA
	36.863684	-76.330088	Norfolk_VA
	35.893286	-75.628164	Roanoke_Island_NC
	31.511252	-81.259024	Sapelo_Island_GA
	31.453722	-81.361311	Sapelo_Island_GA
<i>Mullus barbatus</i>	43.233329	4.166664	Sete
Maggio et al 2009	44.083329	8.783335	Genova
Marine	40.466668	8.133337	Alghero
	38.083332	14.61667	S_Agata_di_Militello
	38.050001	13.58333	Porticello
	38.150001	12.916662	Castellamare_del_Golfo
	37.5	8.666671	Tunisia
	36.816665	13.983331	Sicily
	36.466667	15.299997	Sicily
	37.400001	15.48333	Catania
	44.166666	12.78333	Rimini
	43.899999	13.150004	Fano
	41.333333	17.333336	Bari
	41.816666	18.750004	Albania
<i>Thunnus alalunga</i>	40	1.56667	Med_05
Davis et al 2011	39.816667	12.999997	Med_06
Marine	39.816667	13.000003	Med_07

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
	48.500002	-10.633327	CS1_05
	47.566667	-12.466663	CS2_05
	48.350001	-10.483332	CS3_05
	47.533335	-9.499998	CS1_06
	44.833334	-3.333331	BB1_06
	43.599999	-2.383331	BB2_06
	52.316667	-12.383332	W11_07
	51.55	-13.849997	W12_07
	26.999997	-17.000002	CAN_07
	-21.000001	163.833334	Pac_03
	-14.000003	-134.999999	Pac_05
<i>Leucopsarion petersii</i>	40.785809	140.204048	AM
Kokita, et al 2013	35.714663	136.062993	FK
anadromous	35.016504	138.540124	SZ
	33.608676	135.951343	WK
<i>Acipenser oxyrinchus oxyrinchus</i>	47.044194	-70.706195	St_Lawrence_River
King et al 2001	45.266069	-66.064175	St_John_River
Anadromous	40.733782	-74.01851	Hudson_River
	39.668311	-75.536952	Delaware_R
	36.061922	-76.011682	Albemarle_Sound
	31.324342	-81.322808	Altamaha_River
<i>Retropinna spp.</i>	30.319904	119.424838	TM
Yuan et al. 2014	30.109523	118.884155	LT
Catadromous	29.716173	118.299077	HS
	30.107866	118.863319	QL
	32.817313	131.365723	SB
	32.566737	131.088648	SR
<i>Anguilla anguilla</i>	53.870124	-9.688625	Mayo_Burishoole_River
Daemen et al 2001	43.579087	10.297192	Livorno_Aron_river
Catadromous	34.166664	-6.833332	Kenitra_Sebou_River
	57.201932	12.140582	Viskan_river
	51.687517	-2.543226	Severn_Estuary
<i>Rhinogobius sp.</i>	35.342665	136.031001	ado
Ohara, et al 2009	35.463556	136.076698	L_Biwa
Amphidromous	33.92505	133.129366	Kamo
	35.373079	132.673489	Ino
	35.55342	134.824545	Maruyama
	39.887372	139.944173	Babame
	35.232959	139.868907	Iwase
	35.26656	140.400755	Ochiai
	34.504183	132.892513	Numata
	34.083966	134.578021	Yoshino

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
	35.515368	134.822687	Maruyama
	34.083567	135.11923	Arida
<i>Hypomesus transpacificus</i>	38.053264	-121.960891	Suisum_bay
Fisch et al. 2011	38.180975	-121.996049	Montezuma_Slough
Freshwater	38.087944	-121.747065	Lower_Sacramento_River
	38.253045	-121.687831	Cache_Slough
	38.409534	-121.613609	Deep_Water_Ship_Channel
<i>Neogobius fluviatilis</i>	53.006497	18.611708	A*NF
Neilson & Stepien 2011	54.327894	19.526324	B*NF
Freshwater	47.818608	18.740878	C*NF
	45.391914	29.588268	DNF
	48.566667	26.75	ENF
	46.468333	30.216667	FNF
	46.229242	30.36288	GNF
	46.733333	33.266667	HNF
	46.650683	30.53375	INF
	46.69	31.20345	JNF
	46.655616	35.278634	KNF
	47.114634	40.791515	LNf
	42.135	41.701111	MNF
	46.016147	43.448435	NNF
	48.674471	43.515149	ONF
	48.643269	43.617069	PNF
	45.617691	44.211077	QNF
	47.370632	45.208318	RNF
	46.272008	45.615373	SNF
	48.87087	44.660139	TNF
	48.484638	44.784676	UNF
	48.310313	45.797317	VNF
	47.683923	46.509057	WNF
	47.171956	47.053889	XNF
	46.82965	47.600639	YNF
	46.601411	47.883446	ZNF
	45.78835	47.886953	AANF
	46.302213	48.977384	BBNF
<i>Neogobius pallasii</i>	41.837222	48.62	CCNP
Neilson & Stepien 2011	38.751944	48.868889	DDNP
Freshwater	41.53795	48.924108	EENP
	40.888889	49.368889	FFNP
	39.94	49.409167	GGNP
	40.217222	49.569167	HHNP

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
	40.600278	49.682222	IINP
	40.304167	49.805	JJNP
	40.576667	50.030833	KKNP
<i>Salvelinus alpinus</i>	65.724579	-64.789058	iq2
Harris <i>et al.</i> 2014	66.34629	-64.355012	kin
Anadromous	66.274095	-66.166018	ava
	66.435057	-66.519946	iat
	66.337549	-66.787942	kek
	66.578329	-66.737234	iq2
	66.828495	-68.204556	ius
	66.440033	-67.434357	kan
	66.547613	-67.915551	kip
	65.949506	-67.414139	aun
	65.435318	-67.41068	ikp
	65.236655	-67.258894	opi
	65.04435	-67.11871	iqa
	64.616303	-66.30833	qas
<i>Siphateles bicolor ssp.</i>	38.675155	-119.538737	tpz
Finger & May 2015	39.514724	-118.883229	lsl
Freshwater	39.107011	-119.908186	spn
	38.935527	-120.014435	tkS
	40.007951	-119.561764	pyr
	38.156681	-119.345889	twN
	38.773562	-118.735742	wlk
	39.685408	-118.127921	dxv
	38.433747	-119.015418	ewr
	40.671998	-115.761713	sfh
<i>Oncorhynchus clarki</i>	48.946816	-122.474476	double_ditch
wenburger <i>et al.</i> 1998	48.558942	-122.022139	parker_creek
Freshwater	47.326178	-122.028269	Covington_creek
	47.155836	-122.216735	fennel_creek
	47.559632	-122.832577	Gold_Creek
	47.625056	-122.872888	Stavis_Creek
	48.11551	-123.06813	Gierin_Creek
	48.1062	-123.426893	peabody_creek
	48.155798	-123.704195	salt_creek
	47.832432	-124.497636	goodman_creek
	47.654837	-124.180935	snahapish_river
	47.020212	-123.362107	wildcat_creek
	46.580443	-123.634329	oxbow_creek
<i>Negaprion brevirostris</i>	-3.8492	-33.817422	Atol das Rocas
Feldheim <i>et al.</i> 2001	25.854833	-81.629841	Gullivan

Table A2.1. Continued.

Species / Study / LH	Latitude	Longitude	Population ID
Marine	24.578487	-82.097986	Marquesas
	25.718293	-79.303236	Bimini
	56.209015	162.524804	Russia
	63.487908	-159.492548	yukon
Syngnathus floridae	37.31697	-76.292729	VA
Mobley <i>et al.</i> 2010	34.641021	-76.732299	NC
Marine	27.584546	-82.707911	TB
	29.695146	-85.393458	SJ
	27.836863	-97.0439	TX

Table A2.2

Heterozygosity phylogenetic linear regressions.

	Model	Pagel's λ	AIC	Δ AIC
Phylogenetic linear model	He~ max length	2.38E-09	-57.35677	0
	He~ LH + max length			
	He~ LH	2.61E-09	-55.55821	1.79856
	He~ max length *	2.76E-09	-55.3143	2.04247
	LH	2.75E-09	-49.6462	7.71057
Star Tree Phylogenetic linear model	He~ max length		-57.35677	0
	He~ LH + max length			
	He~ LH		-55.55821	1.79856
	He~ max length *		-55.3143	2.04247
	LH		-49.6462	7.71057
Linear model	He~ max length		-59.35677	0
	He~ LH + max length			
	He~ LH		-57.55821	1.79856
	He~ max length *		-57.3143	2.04247
	LH		-51.6462	7.71057

Table A2.3
Euclidian and shortest path distances

Species	Euclidian	SPD	Species	Euclidian	SPD
Marine			Anadromous		
<i>S. leptorhynchus</i>	1.26E-02	1.35E-05	<i>L. fluviatilis</i>	-3.89E-06	-1.45E-05
<i>T. delaisi</i>	6.09E-05	2.34E-05	<i>L. sealii</i>	1.11E-04	6.06E-05
<i>T. chalcogramma</i>	1.68E-07	1.74E-07	<i>C. nasus</i>	2.36E-05	6.13E-05
<i>A. bicinctus</i>	1.28E-05	1.18E-05	<i>L. petersii</i>	6.85E-09	3.42E-05
<i>E. newberryi</i>	2.96E-03	3.56E-03	<i>A. oxyrinchus</i>	4.89E-05	3.36E-05
<i>C. harengus</i>	1.66E-06	1.56E-06	<i>S. alpinus</i>	7.86E-05	3.58E-05
<i>F. heteroclitus</i>	1.72E-04	1.00E-04	<i>A. pseudoharengus</i>	7.75E-05	5.76E-05
<i>M. barbatus</i>	8.70E-07	-8.03E-08	<i>A. aestivalis</i>	2.04E-06	3.95E-05
<i>T. alalunga</i>	2.63E-06	3.00E-06	Amphidromous		
<i>N. brevirostis</i>	2.78E-06	2.12E-06	<i>S. stimpsoni</i>	-1.51E-07	4.51E-07
Catadromous			<i>Rhinogobius sp.</i>	-3.73E-05	-1.19E-04
<i>L. calcarifer</i>	9.92E-04	2.82E-05	Freshwater		
<i>K. rupestri</i>	1.31E-05	1.14E-05	<i>L. planeri</i>	-1.32E-04	1.25E-05
<i>A. japonicus</i>	1.09E-06	8.97E-07	<i>H. transpacificus</i>	-2.94E-04	-1.53E-04
<i>M. colonorum</i>	6.43E-07	2.74E-05	<i>N. pallasii</i>	-2.01E-04	-1.66E-04
<i>P. urvillii</i>	2.54E-05	2.06E-05	<i>N. fluviatilis</i>	5.10E-04	4.84E-05
<i>A. anguilla</i>	3.04E-06	-1.22E-06	<i>S. bicolor ssp.</i>	-3.09E-06	7.66E-05

Slopes of Isolation by distance for both euclidian and shortest path distances

CHAPTER III
THE EFFECTS OF AN AMPHIBIOUS LIFE HISTORY ON THE
POPULATION STRUCTURE OF A MUDSKIPPER
(*PERIOPHTHALMUS MODESTUS*) IN EAST ASIA

Co-authors on this chapter include Jen-Chieh Shiao, Todd W. Pierson, Jie Zhang, Benjamin M. Fitzpatrick. The Chapter was written by JBC. JBC and BMF designed the study, JBC, JCS, JZ conducted field work. JBC and TWP conducted laboratory work, JBC, TWP and BMF analyzed the data. JBC, JCS and JZ contributed funding to the Study. All co-authors assisted with revisions and writing the manuscript.

Abstract

Amphibious fishes require both aquatic and terrestrial portions of their life cycle. While many fishes can persist on land for a limited period of time, few species have an obligate terrestrial life history stage. Many species of mudskippers in the family *Gobiidae* (subfamily: *Oxudercinae*) are obligate amphibious. The obligatory terrestrial phase comes with an environmental restriction to mudflat habitats during breeding when adults build burrows in the mud to lay their eggs. Some of these fishes tend to spend the majority of their time out of water. Effects of such an environmentally restricted out-of-water phase should be reflected in the population structure of the species. To test whether this is the case, we examined the shuttles hoppfish (*Periophthalmus modestus*) with respect to its population structure throughout its range in the East and South China Seas. Over 250 individuals were collected from ten populations across mainland China, Hainan island, Taiwan, Okinawa and central Japan. We used a targeted capture sequencing method, RADcap, to detect SNPs from 265 loci from each individual. We found that, based on observed levels of genetic variation and population structure, these fish are connected across large areas of continuous coastal areas and across small spans of water. At the same time, however, restricted long-distance dispersal across water has resulted in fragmentation of the species' population structure. These results are consistent with the hypothesis that a limited dispersal phase significantly affects population structure in fishes.

Introduction

Life histories that restrict individuals to particular habitats during specific life stages can have a drastic effect on the structure of those species' populations, and in turn their patterns of diversification (McCulloch *et al.* 2009; McDowall 2001; Corush Ch1; Corush Ch2). Limiting dispersal during larval stages or restricting movement during the adult stage can result in a more fragmented population compared to species that are free to disperse (Waters *et al.* 2000; McDowall 2001; Bracken *et al.* 2015). Population connectivity can be restricted in diadromous fishes that must migrate to marine or freshwater environments for reproduction, in some winged insects with limited ability to fly cannot move between isolated habitats, and in amphibious species, restrictions to water or land (McDowall 1988; McCulloch *et al.* 2009).

In fishes, one life history that limits the movement of individuals is the amphibious life history, in which individuals spend time out of water as normal parts of their life history (Gordon *et al.* 1969). While this is rare in fishes, it has evolved repeatedly and is found in well over 100 species (Sayer and Davenport 1991; Martin 2015; Turko and Wright 2015; Wright and Turko 2016). Many fishes including eels, gobies, killifishes, snakeheads, and lungfish have adapted to being out of water (Taylor *et al.* 2007; Sayer 2005; Ishimatsu 2012; Martin 2015). Multiple reasons for emersion have been proposed and tested including poor water quality, avoiding predation, reproduction, and obtaining resources (Sayer and Davenport 1991; Sayer 2005). These imply that a fish, over its life time, may choose to stay in a less optimal aquatic environment and leave the water on occasion as opposed to moving to a better environment. In the many obligate amphibious species, habitat selection is dependent on both the aquatic and terrestrial environment, putting a number of restrictions on where individuals can move. Mudskippers of the genus *Periophthalmus* are just one of many groups of fishes that spawn on beaches (Martin 2015). The mudskipper *Periophthalmus modestus* lays eggs that require air for proper development (Ishimatsu *et al.* 2007). This obligatory phase comes with a restriction to mangrove and mudflat habitats during breeding. While they do not have obligate juvenile or adult stages out of water, these particular fishes tend to spend the majority of their time out of the water (Sayer and Davenport 1991).

The mudskipper, shuttles hopfish, *Periophthalmus modestus* (Cantor 1842), has a distribution across the South and East China seas (China Seas). Individuals spend about 50% of their time out of water and can survive up to 30 hours out of water (Sayer and Davenport 1991). While closely related species such as *P. argentilineatus* are commonly found across mudflats and mangroves, *P. modestus* is typically more geographically restricted to mudflats (Oshiro *et al.* 2005). They build "J" shaped burrows in which the eggs attached to the upper layer of the chamber and are exposed to air for several hours, followed by immersion in water to hatch (Ishimatsu *et al.* 1998, 2007, 2009; Ishimatsu and Gonzales 2011). While the trait has not been explicitly tested in *P. modestus*, eggs of a close relative, *P. argentilineatus*, will die if there is no emersion period (Brillet 1976). To provide this biphasic environmental shift, the male of the species carries air in his

cheeks down the long arm of the “J” and releases it into the small dome-like area deep in the mud. This is timed with tidal cycles so the nest will be flooded and the eggs will be submerged at the proper time during their development. Eggs completely submerged in water or those that did not return to water before the hatching window do not survive (Ishimatsu *et al.* 2007). This complex biphasic lifestyle limits *Periophthalmus* species to a specific environment during reproduction. The question arises: Is this limitation drastic enough to affect population structure?

Movement around the South and East China seas is regulated by currents plus geologic and historic process that have shaped the landscape across the species’ range. A contemporary factor shaping the biotic and abiotic patterns of the China seas is the Kuroshio current. The Kuroshio starts where the North Equatorial Current runs into the Philippines and flows north on the east side of Taiwan, the west side of the Ryukyu Islands over the Okinawa Trough, and then bifurcates to the Tsushima current, which flows north along the western side of the main Japanese islands into the Sea of Japan and to the east side of the main Japanese islands until it feeds into the North Pacific Current. The Kuroshio current plays a major role in the physical features of the water. Both temperature and salinity increase near the edge of the continental shelf as water enters the Kuroshio current system, though the shelf water does not make it through the Kuroshio resulting in a regular variation in water conditions across the region (Nozaki *et al.* 1989). Biogeographical patterns of many species are a function of this system (Gallagher *et al.* 2015). While the Kuroshio current follows a regular path, the water between the Kuroshio current and China, a different set of more variable currents, change seasonally (Yanagi *et al.* 1993). In the winter the China Sea Current brings cooler, less saline water south through the Taiwan Strait, and in the summer, the South China Sea Surface Current carries slightly warmer water and saline water north (Jan *et al.* 2002).

In addition to the currents, there is a large amount of variation in bathymetry within the region. Most of the water between mainland China and Hainan island as well as Taiwan is less than 200M deep, however the Okinawa Trough goes down more than 2500M at its deepest point, separating China and Taiwan from many parts of Japan by deep water (Figure 3.1). Gene flow in many marine organisms is restricted by bathymetric constraints (Schultz *et al.* 2008; Knutsen *et al.* 2009)

Variation on a geological time scale has also altered the dynamics of the China Seas. The islands inhabited by *P. modestus* are comprised both continental islands (e.g. Hainan, Taiwan, Japan) as well as oceanic islands (e.g. Ryukyu islands). While contemporary movement to and from any island would require dispersal, past geological suggest that vicariance could be responsible for movement between mainland Asia and continental islands. lower sea levels associated with past geologic events resulted in continuous coast line from mainland Asia to Hainan, Taiwan, and Japan, but not Okinawa. Taiwan was connected to mainland Asia up until around 11,000 years ago (Voris 2000) and Hainan island even more recently.

An early study using allozymes showed very low levels of molecular variation around the island of Taiwan in *P. modestus* (*P. cantonensis* = jun. syn. of *P. modestus*) (Chang and Lee 1994). Concentrating on Japan, Mukai, and Sugimoto (2006) used mitochondrial DNA NADH dehydrogenase subunit 5 (ND5) gene to show the population on Okinawa was differentiated from the other six populations from the major Japanese islands. A more geographically extensive study added five populations from China to the initial seven sampled in Mukai and Sugimoto (2006) to reassess the genetic patterns of ND5 (He *et al.* 2015). Based on the more extensive analysis, differentiation between populations in the East China Sea and those in the South China Sea was found. Additionally, differentiation between populations in Japan and those in the East China Sea were associated with the Kuroshio current.

When one assesses the population structure of *Periophthalmus modestus*, multiple factors come into play. The species' biology limits its ability to move across vast areas of deep water that do not offer refuge from the water. Migration between locations is thus potentially limited by access to land. Historically, the China Seas region have very different bathymetry and Hainan, Taiwan and Japan were all connected to mainland Asia. Movement across this range, without leaving the coastline, would have been possible in the past. However, oceanic island such as Okinawa have never been connected to mainland Asia and would require dispersal across deep waters of the Kuroshio current.

The amphibious life history limits the movement of adult individuals, should result in increased populations structure when water acts as a barrier between populations. Here we incorporate population genomics with landscape data and demographic models to understand what factors are important in structuring the populations of *P. modestus*. We test that hypothesis that water acts as a barrier for this fish. In particular, we test that deep waters, such as the Okinawa Trough which separates oceanic islands, result in isolated populations, compared to continental islands, which have allowed for formation of land bridges during low sea levels, as well as shallower water between contemporary landmasses.

Methods

Sample collection

A total of 236 samples of *Periophthalmus modestus* were collected between 2014 and 2017 from ten locations across the South and East China Seas including two populations from Taiwan (Taoyuan and Taichung), five population from China (Cixi, Hainan, Xiapu, Xiashan, and Zhoushan), two populations on mainland Japan, hereafter referred to as “Japan”, (Ariake and the Seto Sea), and one populations from Okinawa, Japan (Figure 3.1). All animals were collected with dip nets, and tissue samples were taken and immediately stored in 95% molecular grade EtOH. Fishes were then fixed in formalin or 70% EtOH. All samples were collected in accordance with protocols approved by University of Tennessee-Knoxville IACUC (UTK#2548-0616). All samples were collected in accordance with each country's regulations.

3RAD DNA Sequencing for RADcap Bait Design

Genomic DNA was extracted using DNeasy Blood and Tissue kits (Qiagen). Initial sequencing was done with a representative subset containing samples from each population to identify target loci for full analysis. Initial sampling was done using the three enzyme RADseq method (Glenn *et al.* 2017). The digestion step included 1.5 μ L 10x CutSmart® Buffer, 4.5 μ L of dH₂O, 0.5 μ L of MspI at 20 U/ μ L, 1.0 μ L of ClaI at 10 U/ μ L, 0.5 μ L of BamHI-HF at 20 U/ μ L, 1 μ L 5 μ M double-stranded iTru R1.A adapter, 1 μ L 5 μ M double-stranded iTru R2.1 adapter, and 5 μ L DNA. Restriction enzymes and CutSmart® Buffer were products of New England Biolabs (NEB) (Beverly, 154 MA, USA). Samples were digested for 1 hr at 37 °C. Following digestion ligation was completed by adding ligation mix to each sample. Ligation mix consisted of 2.0 μ L dH₂O, 1.5 μ L ATP (10 μ M), 0.5 μ L 10x Ligase Buffer, 1.0 μ L T4 DNA Ligase (100 units/ μ L, NEB M0202L buffer diluted 1:3 in NEB B8001S 255 enzyme dilution buffer). Ligation conditions were two cycles of 22°C (20 min) and 37°C (10 min) followed by 80°C (20 min). Following the ligation, 30 μ L dH₂O was added to each 20 μ L ligation product. 1 μ L 5 μ M double-stranded iTru R1.A adapter, 1 μ L 5 μ M double-stranded iTru R2.1 adapter, and 5 μ L DNA. Restriction enzymes and CutSmart® Buffer were products of New England Biolabs (NEB) (Beverly, 154 MA, USA). Samples were digested for 1 hr at 37 °C. Following digestion ligation was completed by adding ligation mix to each sample. Ligation mix consisted of 2.0 μ L dH₂O, 1.5 μ L ATP (10 μ M), 0.5 μ L 10x Ligase Buffer, 1.0 μ L T4 DNA Ligase (100 units/ μ L, NEB M0202L buffer diluted 1:3 in NEB B8001S 255 enzyme dilution buffer). Ligation conditions were two cycles of 22°C (20 min) and 37°C (10 min) followed by 80°C (20 min). Following the ligation, 30 μ L dH₂O was added to each 20 μ L ligation product. NaCl-PEG diluted SpeedBeads (Rohland and Reich 2012) were used for bead cleanup to remove all added enzymes from DNA. Beads were added at a 1.2:1 SpeedBead to ligation product ratio. Bead cleanup procedure followed Glenn *et al.* (2017). Bead cleanup product was resuspended to 25 μ L. Cleaned product was then used for PCR containing. 5.0 μ L 5X Kapa HiFi Buffer (Kapa Biosystems, Wilmington, MA), 0.75 μ L dNTPs (10 μ M), 3.75 μ L dH₂O, 0.5 μ L Kapa HiFi DNA Polymerase (1 unit/ μ L), 2.5 μ L iTru5 primer (5 μ M), 2.5 μ L iTru7 primer (5 μ M) and 10.0 μ L cleaned up DNA. PCR parameters were: 95 °C (2 min) initial denaturing followed by 14 cycles of 98°C (20 sec), 60°C (15 sec) and 72°C (30 sec), with a final extension at 72°C (5 min) followed by a hold at 15°C. Following PCR, a second round of bead cleanup was performed adding 100 μ L of sample to 83 μ L bead mix. Cleaned up product was resuspended to 30 μ L. Concentration of cleaned PCR product was determined using a Qubit Fluorimeter (Life Technologies, Inc.) and all 13 samples were pooled according to concentrations. For size selection, 30 μ L of cleaned pooled libraries was then size selected for 550 +/- 10% bp using Pippin Prep™ (Sage Science, Inc.) following the suggested protocol. Samples were then pooled with other projects and sequenced on an Illumina NextSeq v2 300 cycle kit to obtain 150 bp paired-end (PE150) reads at the Georgia Genomics Facility.

Baits design

Illumina data files were first demultiplexed (see Additional File A3.1 for ipyrad parameters file). ipyrad v0.7.17 (<http://ipyrad.readthedocs.io/>) was used to process reads and assemble loci. Sequences were then aligned to *Periophthalmus magnuspinnatus* genome (You *et al.* 2014; GenBank#: JACL000000000.1). After sequences were aligned with the *P. magnuspinnatus* genome, 2331 loci were identified that had at least 250 bp. Samples with four or more SNPs were removed leaving 1295 loci with zero to three SNPs and at least 250bp long. Of the 1295 loci, 1000 were chosen randomly and used for RADcap sequencing.

RADcap

Genomic DNA from all 236 samples underwent the digestion and ligation step as in the 3RAD methods (above) with the following changes. Digestion mix contained 10 µL of DNA and 0 µL dH₂O. Ligation mix contained 2.75 µL dH₂O and only 0.25 µL of T4 DNA Ligase. Samples were quantified and pooled proportionately. A maximum of 96 samples were combined in each pool. These pools were split into two tubes and cleaned with a 1.2:1 ratio of SpeedBeads to pooled ligation products. Cleaned product was reconstituted in 33 µL dH₂O and each of the two replicates were further split in two (i.e., a total of four tubes per pool). A single cycle of PCR was then performed with 10.0 µL 5X Kapa HiFi Buffer (Kapa Biosystems, Wilmington, MA), 1.5 µL dNTPs (10 266 µM), 1.0 µL Kapa HiFi DNA Polymerase (1 unit/µL), 5.0 µL 8N primer and 15.0 µL DNA. PCR parameters were: 98 °C (1 min), 60°C (30 sec) and 72°C (6 min) followed by a hold at 12°C. Following this PCR, the four replicates were pooled, cleaned again using a 1.2:1 SpeedBead to PCR product ratio, and reconstituted in 30 µL dH₂O. We then conducted a limited-cycle PCR with three replicates per pool with 10 µL DNA, 5 µL P5 primer, 5 µL iTru7 primer, 10 µL buffer, 1.5 µL dNTPs, 1 µL polymerase, and 17.5 µL dH₂O. We then pooled these replicates, cleaned them with a 2:1 SpeedBead to PCR product ratio, and reconstituted them in 40 µL dH₂O. Samples were quantified. Before going into the MYbaits protocol, we conducted a test PCR step to verify all previous steps. The PCR consisted of 1 µL DNA, 2.5 µL P5 primer, 2.5 µL P7 primer, 5 µL buffer, 0.75 µL dNTPs, 12.75 µL dH₂O, and 0.5 µL polymerase and followed the same protocol and for 15 cycles. We verified this product on a 1% agarose gel.

MYbaits® Capture Protocol and Sequencing

Pooled RADcap product underwent in-solution sequence capture according to the MYbaits® protocol (version 3.02, July 2016). Hybridization temperature was set to 65°C. The library elution and amplification's extension time was 1:00.

Bioinformatics

Illumina data files were first demultiplexed based on external tags using BCL2FASTQ resulting in a single file for each individual MYbaits® Capture. Stacks (Catchen *et al.* 2013) was then used to remove PCR duplicates based on the 8N tags added in the RADcap procedure after ligation. We used ipyrad v0.7.17 (<http://ipyrad.readthedocs.io/>)

to process reads, map them against a reference genome we created during the baits design (see Additional File A3.2 for ipyrad parameters file).

Genetic Analysis

Population structure was assessed using STRUCTURE version 2.3.4 (Pritchard 2000). All STRUCTURE runs were for 500,000 generations with a burn in of 100,000 generations. Each analysis was run three times. The complete analysis ran at K=1-9. Based on the results, further Subpopulation STRUCTURE analysis for the “China/Taiwan” (ran K=1-6) and the “Okinawa/Japan cluster” (ran K=1-4) were both run separately. STRUCTURE HARVESTER (Earl 2012) was used to determine optimal number of clusters (K) from STRUCTURE analyses based on likelihood.

Output files from ipyrad were read into R (R core team 2013) using the datSTR() function in the adagenet package (Jombart 2008). Data were then converted from genind format to a usable data frame format using the genind2hierfstat() function in the hierfstat package (Goudet 2005). Pairwise F_{st} values (Weir and Cockerham 1984) were calculated using the pairwise.WCfst() functions and mean expected heterozygosity was calculated overall and for each populations using the basic.stats() function, both in hierfstat. Based on structure analysis, pairwise F_{st} was also run using pooled populations of 1) China, 2) Taiwan, 3) Okinawa, 4) mainland Japan.

Geographic Analyses

Pairwise Euclidian distances (EUC) between each population were calculated in ArcMap using the Euclidian distance tool in the special analysis toolkit. Least cost path distance was then calculated based on bathymetric data obtained from the Natural Earth database (<http://www.naturalearthdata.com>) (Becker *et al.* 2009) using the “Least-Cost Corridors and Least-Cost Paths: Pairwise Comparison” tool in SDMtoolbox 2.2c (Brown 2014; Brown *et al.* 2017). Based on available data, ocean depth layers used in the least cost path analysis included 200M (Continental shelf depth), 1000M, 2000M, 3000M, 4000M, 5000M and 6000M. This provided both a least cost path distance (LCPD (the length of the path of least resistance) as well as a least cost path (LDP) (path of least resistance with higher resistance for deeper water). Variatus permutations of mantel and partial mantel test was preformed to compare pairwise Genetic variation ($F_{st}/1 - F_{st}$) values with EUC, LCP, and/or LCPD using the mantel() and mantel.partial() functions with method = "pearson" in the vegan package (Oksanen *et al.* 2013). Genetic variation ($F_{st}/1 - F_{st}$) was also regressed against various metrics of geographic distance (EUC, LCP and LCPD) to visualize data.

Additionally, migration likelihood surface was created using EEMS (Petkova *et al.* 2016) which estimates migration rates based on genetic dissimilarity. We used an estimate of the species range, based on available records (fishbase and Fishnet2) and ran analyses using 200, 400, and 600 demes. All MCMC samplers were ran for 20,000,000 generations.

Demographic history

Coalescence analyses were run using strataG (Archer *et al.* 2017) a wrapper for Fastsimcoal2 (Excoffier *et al.* 2013). Multiple scenarios were run to test vicariance vs. dispersal between China, Taiwan, Okinawa and Japan. Analyses allowed for various patterns of colonization and migration between the four populations (Additional file A3.3-A3.7). Models included an equilibrium island model (IM), two stepping stone models with Japan and Taiwan being colonized from China and Okinawa being colonized from Taiwan (model SST) or Okinawa being colonized from Japan (model SSJ). Lastly, two vicariance/colonization models were simulated where all populations were colonized from China, with one model simulating exponential expansion of Okinawa (model VCe) vs. consistent population size in all populations (model VC). All scenarios were simulated 20,000 times. Summary statistics were calculated from each simulation including pairwise F_{st} , Global F_{st} , allelic richness and population heterozygosity. Model selection with Approximate Bayesian Computation was conducted using the ABC package (Csilléry *et al.* 2012) based on 10,000 simulations. For each scenario the posterior probabilities were compared using the multinomial logistic regression method in the postpr() function. Comparison of scenario was based on the proportion of accepted simulations given this model. For each scenario population sizes, migration rates, mutation rates and colonization event times were estimated from various distributions (See Additional File A3.3- A3.7 for parameters). Parameter estimates based on 200 posterior samples from 20,000 runs. Sample size remained consistent with observed data: China= 119, Japan=55, Taiwan=50 and Okinawa=12.

Results

Baits design and Sequencing

1000 loci were identified and 2 baits per locus were designed. All loci were verified to be compatible by MYcroarray as part of the MYbaits® capture design. Of the 1000 targeted loci, 365 were consistently sequenced across samples with high coverage. Filtering resulted in a total 365 loci for 236 individuals. Mean sample coverage was 353 loci (Table A3.1 for full distribution)

Genetic Analysis

Overall heterozygosity (H_e) was 0.0343 with a range of 0.01 (Okinawa)-0.046 (Ariake, Japan) and a mean of 0.338 (Table 3.1). With the exception of Okinawa and Ariake, all populations had fairly similar levels of heterozygosity. The pairwise F_{ST} values were typically largest between Okinawa and all other locations. All comparisons involving Okinawa were over 0.26 whereas all values not including Okinawa were less than 0.17 (Table 2).

STRUCTURE and STRUCTURE HARVESTER analysis indicate that the main division in populations is across the East China Sea with Japan, and Okinawa forming one group and Mainland China, Hainan and Taiwan forming another group. Subgroup STRUCTURE analyses further indicate the genetic makeup of *P. modestus* on Mainland

China and Hainan differs from Taiwan and that Japan and Okinawa house distinct populations. (Figure 3.2)

Geographic Analysis

EEMS migration analysis showed, when more demes were used, migration rates are high along the coast of China and around Taiwan, but low in the East China Sea, particularly between Taiwan and China and around Okinawa (Figure 3.3).

Least cost path seemed to go around the Okinawa trough when populations were on different sides of the East China Sea (Figure 3.1). This along with the cost of dispersal in deep water change the corrected distance of travel by incorporating both the LCP and LCPD (Figure 3.4). Models incorporating LCP and LCPD were better fit then just Euclidian distance. (Table 3.3)

Demographic history

The optimal scenario (VCe) had 83.26% accepted simulations. The next best scenario had an acceptance rate of 7.22% (Table 3.4). The optimal scenario indicates that populations on Taiwan, Japan and Okinawa all arose from those on China. Migrations rate estimations between all populations were low with each having a mean value of ≤ 0.0003 . The estimated time of coalescence between China and Taiwan, Japan and Okinawa were 2890, 2924, and 2857 generations respectively. Estimated populations sizes for China, Taiwan, Japan, and Okinawa were 73224, 1675, 5666, and 5372 respectively (Table 3.5).

Discussion

Based on analyses of genetic variation and structure in *Periophthalmus modestus*, the fish are connected across large areas of continuous coastal areas (e.g., mainland China) but restricted in movement by even small spans of shallow water (e.g., Taiwan strait, Qiongzhou Strait, and the multiple straits between Japan's main island Honshu, Shikoku, and Kyushu), as well as between long-distance open water from Japan to China and from Okinawa to China, Japan, and Taiwan. These restrictions result in fragmentation of the species' population structure. These results are consistent with the hypothesis that a limited dispersal ability, as a result of obligate amphibious behavior, may significantly affect population structure in fishes. We propose that populations on continental island were established as a result of vicariance from mainland Asia when sea levels rose and isolated Taiwan, Japan, and Hainan islands. Around this time, when mainland Asia was closer to Okinawa, colonization from mainland Asia to Okinawa would have required a shorter distance across shallower waters compared to today's geology. Population structure and migration analyses support movement between the East and South China Seas along the coast of China, but lower migration rates between any populations separated by water.

Fastsimcoal2 results suggest that Taiwan, Japan, and Okinawa were all colonized from China before sea levels rose and disconnected Taiwan and Japan from mainland Asia. In

fact, many examples that support vicariance within the region (Shih *et al.* 2007). The model also suggests that the population on Okinawa originated during or before sea levels were lower. In addition, simulated migration rates are extremely small indication little migration between China, Taiwan, Okinawa, and Japan. This pattern would explain the low heterozygosity and high population differentiation between Okinawa and other locations. Additionally, since Taiwan and Japan have been isolated, low levels of migration have been occurring across water. The higher levels of heterozygosity and lower levels of population differentiation must be the result of one continuous population along the coast of East Asia that was split up with lowering of sea levels.

These results agree with those of previous studies that show a strong level of isolation between Okinawa and other populations (Mukai and Sugimoto 2006). The question arises, how does an obligate and highly terrestrial fish get to Okinawa? Although very small migration rates are inferred, multiple individuals have to have colonized the island at least once. In order to reach Okinawa, the fishes must then have dispersed somehow. Other amphibious fishes, killifishes, have been found in dead tree trunks, which could act as a mode of transportation (Taylor *et al.* 2007). Mudskipper behavior is not similar to that of killifishes and mudskippers would be unlikely to hide in dead logs for extended periods. However, in a region prone to seasonal storms, it is quite possible that vegetation carrying hitchhikers has moved between islands. Additionally, past sea levels decreased the distance between Okinawa and mainland Asia, allowing for a shorter dispersal between the landmasses.

These results also suggest that isolation by distance in aquatic organisms is a complex concept. Traditional methods of assessing a species' populations by looking at genetic compared to geographic distance, even when incorporating biologically meaningful resistance layers, may show patterns that do not represent traditional isolation by distance. This data shows three groupings of points, driven by a few isolated (Okinawa) and many connected (mainland China) populations (figure 3.4). However, isolation by distance slopes are not representative of the overall level of isolation across distances, but instead driven by the large variation in genetic differentiation between a few pairwise comparisons. Therefore, we suggest that more in-depth information about the natural history, geographic history, and evolutionary history of species should be incorporated into population genetics studies, as we attempt to do here. Demographic models and multiple analyses of population structure and differentiation are necessary for a better understanding of what shapes a species' movement.

Additional sampling along the Ryukyu island chain could shed light on how dispersal between oceanic island occurs. Additional sampling along the Ryukyu islands and South Korea would be particularly informative in allowing a better understanding the dynamics of *Periophthalmus modestus*' biology. Additionally, comparison between less restricted amphibious fishes such as *P. argentilineatus* (which is found across extremely isolated island such as Guam and Palau) or less terrestrial species such as those in the genus *Scartelaos* would be informative of how restrictive amphibious behavior really is.

Acknowledgments:

We would like to thank Travis C Glenn, Natalia J Bayona-Vasquez, Troy J Kieran for their assistance with the 3RAD and RADcap methods. Field work would not have been possible without the help of Mr. Feng Lin, from the Key Laboratory of Animal Ecology and Conservation Biology Institute of Zoology, Chinese Academy of Sciences, We would also like to thank Graciela Cabana, Ben Keck, and Danial Simberloff for their comments throughout the project. This research was funded by NSF-EAPSI 2014 (Award#1414902), NSP-EAPSI 2016 (Award#1614100), National Natural Science Foundation of China #31272287, and the University of Tennessee Department of Ecology and Evolutionary Biology.

References

- Archer, F.I., Adams, P.E. and Schneiders, B.B., 2017. stratag: An r package for manipulating, summarizing and analysing population genetic data. *Molecular ecology resources*, 17(1), pp.5-11.
- Becker, J.J., Sandwell, D.T., Smith, W.H.F., Braud, J., Binder, B., Depner, J., Fabre, D., Factor, J., Ingalls, S., Kim, S.H. and Ladner, R., 2009. Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30_PLUS. *Marine Geodesy*, 32(4), pp.355-371.
- Bracken, F.S., Hoelzel, A.R., Hume, J.B. and Lucas, M.C., 2015. Contrasting population genetic structure among freshwater-resident and anadromous lampreys: the role of demographic history, differential dispersal and anthropogenic barriers to movement. *Molecular ecology*, 24(6), pp.1188-1204.
- Brillet C (1976) Structure du terrier, reproduction et comportement des jeunes chez le poisson amphibie *Periophthalmus sobrinus* Eggert. *Rev Ecol Terre Vie* 30:465–483
- Brown JL, Bennett JR, French CM (2017). SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ*
- Brown J.L. 2014, SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic, and species distribution model analyses. *Methods in Ecology and Evolution* DOI: 10.1111/2041-210X.12200
- Catchen, J., P. Hohenlohe, S. Bassham, A. Amores, and W. Cresko. Stacks: an analysis tool set for population genomics. *Molecular Ecology*. 2013.
- Catchen, J., A. Amores, P. Hohenlohe, W. Cresko, and J. Postlethwait. Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, 1:171-182, 2011.
- Chang, J.T. and Lee, S.C., 1994. Genetic variation of the Chinese mudskipper, *Periophthalmus cantonensis* (Osbeck, 1762) (Pisces; Perciformes, Gobiidae) from Taiwan. *Zool. Stud*, 33(1), pp.34-43.
- Csilléry, K., François, O. and Blum, M.G., 2012. abc: an R package for approximate Bayesian computation (ABC). *Methods in ecology and evolution*, 3(3), pp.475-479.
- Earl, D.A., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources*, 4(2), pp.359-361.

- Eaton, D.A., 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, 30(13), pp.1844-1849.
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C., and M. Foll (2013) Robust demographic inference from genomic and SNP data. *PLOS Genetics*, 9(10):e1003905
- Gallagher, S.J., Kitamura, A., Iryu, Y., Itaki, T., Koizumi, I. and Hoiles, P.W., 2015. The Pliocene to recent history of the Kuroshio and Tsushima Currents: a multi-proxy approach. *Progress in Earth and Planetary Science*, 2(1), p.17.
- Glenn, T.C., Bayona-Vasquez, N.J., Kieran, T.J., Pierson, T.W., Hoffberg, S.L., Scott, P.A., Bentley, K.E., Finger, J.W., Watson, P.R., Louha, S. and Troendle, N., 2017. Adapterama III: Quadruple-indexed, triple-enzyme RADseq libraries for about \$1 USD per Sample (3RAD). *bioRxiv*, p.205799.
- Goudet, J., 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular ecology resources*, 5(1), pp.184-186.
- He, L., Mukai, T., Chu, K.H., Ma, Q. and Zhang, J., 2015. Biogeographical role of the Kuroshio Current in the amphibious mudskipper *Periophthalmus modestus* indicated by mitochondrial DNA data. *Scientific reports*, 5, p.15645.
- Hoffberg, S.L., Kieran, T.J., Catchen, J.M., Devault, A., Faircloth, B.C., Mauricio, R. and Glenn, T.C., 2016. RADcap: sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Molecular ecology resources*, 16(5), pp.1264-1278.
- Ishimatsu, A., 2012. Evolution of the cardiorespiratory system in air-breathing fishes. *Aqua-BioScience Monographs*, 5(1), pp.1-28.
- Ishimatsu, A. and Gonzales, T.T., 2011. Mudskippers: front runners in the modern invasion of land. *The biology of gobies*. Enfield: Science Publishers, pp.609-38.
- Ishimatsu, A., Hishida, Y., Takita, T., Kanda, T., Oikawa, S., Takeda, T. and Huat, K.K., 1998. Mudskippers store air in their burrows. *Nature*, 391(6664), p.237.
- Ishimatsu, A., Takeda, T., Tsuchiko, Y., Gonzales, T.T. and Khoo, K.H., 2009. Direct evidence for aerial egg deposition in the burrows of the Malaysian mudskipper, *Periophthalmodon schlosseri*. *Ichthyological research*, 56(4), p.417.
- Ishimatsu, A., Yoshida, Y., Itoki, N., Takeda, T., Lee, H.J. and Graham, J.B., 2007. Mudskippers brood their eggs in air but submerge them for hatching. *Journal of Experimental Biology*, 210(22), pp.3946-3954.

Jan, S., Wang, J., Chern, C.S. and Chao, S.Y., 2002. Seasonal variation of the circulation in the Taiwan Strait. *Journal of Marine Systems*, 35(3-4), pp.249-268.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), pp.1403-1405.

Knutsen, H., Jorde, P.E., Sannaes, H., Rus Hoelzel, A., Bergstad, O.A., Stefanni, S., Johansen, T. and Stenseth, N.C., 2009. Bathymetric barriers promoting genetic structure in the deepwater demersal fish tusk (*Brosme brosme*). *Molecular Ecology*, 18(15), pp.3151-3162.

Martin, K.L., 2014. Beach-spawning fishes: reproduction in an endangered ecosystem. CRC Press.

McCulloch, G.A., Wallis, G.P. and Waters, J.M., 2009. Do insects lose flight before they lose their wings? Population genetic structure in subalpine stoneflies. *Molecular ecology*, 18(19), pp.4073-4087.

McDowall, R.M., 1988. *Diadromy in fishes: migrations between freshwater and marine environments*. Croom Helm.

McDowall, R.M., 2001. Diadromy, diversity and divergence: implications for speciation processes in fishes. *Fish and Fisheries*, 2(3), pp.278-285.

Mukai, T. and Sugimoto, M., 2006. Genetic population structure of the mudskipper, *Periophthalmus modestus*, in Japan inferred from mitochondrial DNA sequence variations. *Japanese Journal of Ichthyology*, 53(2), pp.151-158.

Nozaki, Y., Kasemsupaya, V. and Tsubota, H., 1989. Mean residence time of the shelf water in the East China and the Yellow Seas determined by ²²⁸Ra/²²⁶Ra measurements. *Geophysical Research Letters*, 16(11), pp.1297-1300.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H. and Oksanen, M.J., 2013. Package 'vegan'. *Community ecology package, version*, 2(9).

Oshiro, M., Tsubako, Y., Tachihara, K., 2005. Habitat selection for burrow making by two species of mudskippers, *Periophthalmus argentilineatus* and *P. modestus* on the Okinawa island, Japan. *Okinawa Biol. Educ.* 37, 9–18.

Petkova, D., Novembre, J. and Stephens, M., 2016. Visualizing spatial population structure with estimated effective migration surfaces. *Nature genetics*, 48(1), p.94.

- Pritchard, J.K., Stephens, M. and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155(2), pp.945-959.
- Sayer, M.D., 2005. Adaptations of amphibious fish for surviving life out of water. *Fish and Fisheries*, 6(3), pp.186-211.
- Sayer, M.D.J. and Davenport, J., 1991. Amphibious fish: why do they leave water?. *Reviews in Fish Biology and Fisheries*, 1(2), pp.159-181.
- Schultz, J.K., Feldheim, K.A., Gruber, S.H., Ashley, M.V., McGovern, T.M. and Bowen, B.W., 2008. Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). *Molecular Ecology*, 17(24), pp.5336-5348.
- Shih, H.T., Fang, S.H. and Ng, P.K., 2007. Phylogeny of the freshwater crab genus *Somanniathelphusa* Bott (Decapoda: Parathelphusidae) from Taiwan and the coastal regions of China, with notes on their biogeography. *Invertebrate Systematics*, 21(1), pp.29-37.
- Taylor, D.S., Turner, B.J., Davis, W.P. and Chapman, B.B., 2007. A novel terrestrial fish habitat inside emergent logs. *The American Naturalist*, 171(2), pp.263-266.
- Team, R.C., 2013. R: A language and environment for statistical computing.
- Wright, P.A. and Turko, A.J., 2016. Amphibious fishes: evolution and phenotypic plasticity. *Journal of Experimental Biology*, 219(15), pp.2245-2259.
- Turko, A.J. and Wright, P.A., 2015. Evolution, ecology and physiology of amphibious killifishes (Cyprinodontiformes). *Journal of fish biology*, 87(4), pp.815-835.
- Voris, H.K., 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, 27(5), pp.1153-1167.
- Waters, J.M., Dijkstra, L.H. and Wallis, G.P., 2000. Biogeography of a southern hemisphere freshwater fish: how important is marine dispersal?. *Molecular Ecology*, 9(11), pp.1815-1821.
- Weir, B.S. and Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *evolution*, 38(6), pp.1358-1370.
- Yanagi, T. and Takahashi, S., 1993. Seasonal variation of circulations in the East China Sea and the Yellow Sea. *Journal of Oceanography*, 49(5), pp.503-520.

You, X., Bian, C., Zan, Q., Xu, X., Liu, X., Chen, J., Wang, J., Qiu, Y., Li, W., Zhang, X. and Sun, Y., 2014. Mudskipper genomes provide insights into the terrestrial adaptation of amphibious fishes. *Nature communications*, 5, p.5594.

Appendix A – Figures and Tables

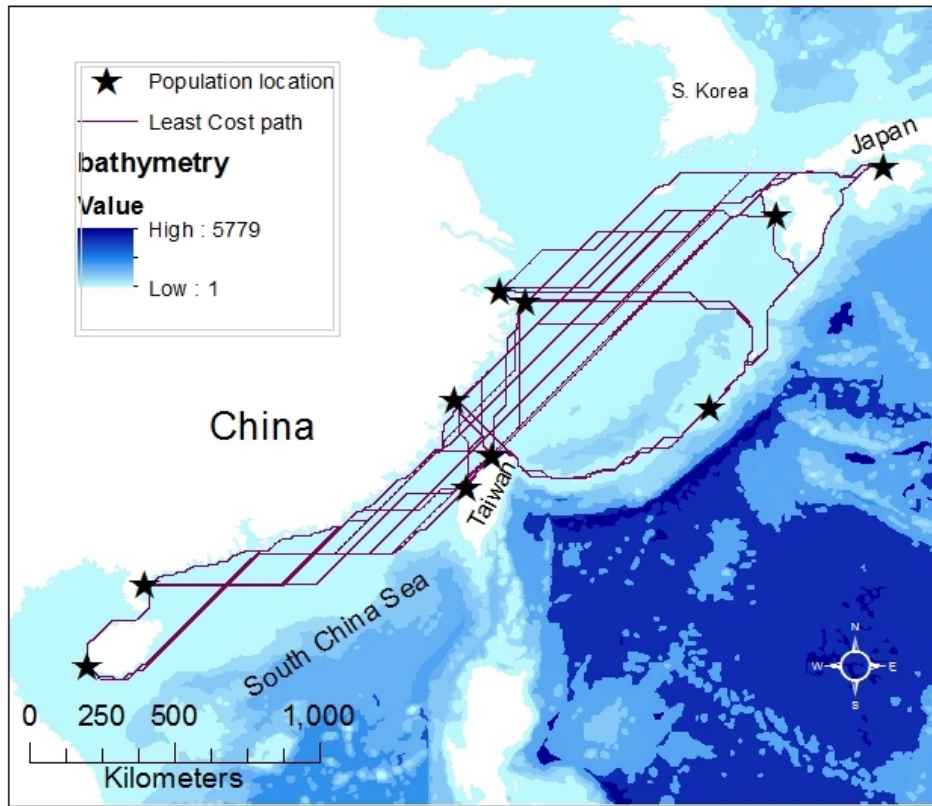
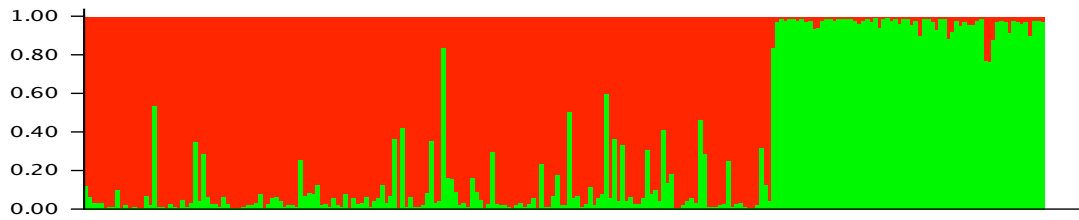


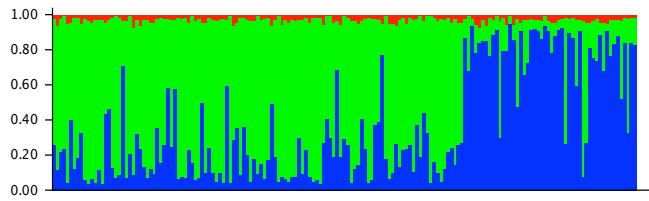
Figure 3.1 Population locations of *Periophthalmus modestus* across the South and East China seas.

Stars represent samples sites of ten populations. Lines represent the least cost paths between each pair of populations based on resistance. Resistance was based on bathymetry values shown in various shades of blue.

A)



B)



C)

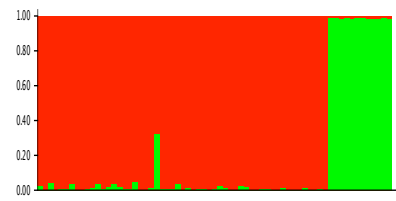


Figure 3.2 STRUCTURE population assignment plots.

All STRUCTURE runs were for 500,000 generations. The complete analysis (run at $K=1-10$) separated Taiwan and China from all of Japan. Here is the complete analysis groups: Japan (Primarily red) from Taiwan/China (Green) b) Subpopulation STRUCTURE analysis: For the “China/Taiwan” cluster the optimal is $K=3$ (ran $K=1-7$), with the two mostly red section being the Taiwan samples (#50-69, 140-169) and the primarily green blocks representing China, (#1-49, 70-139). C) For the “Japan cluster” (ran $K=1-4$) the optimal $K=2$. The red bars are the Okinawa samples and the Green are the two Japan populations.

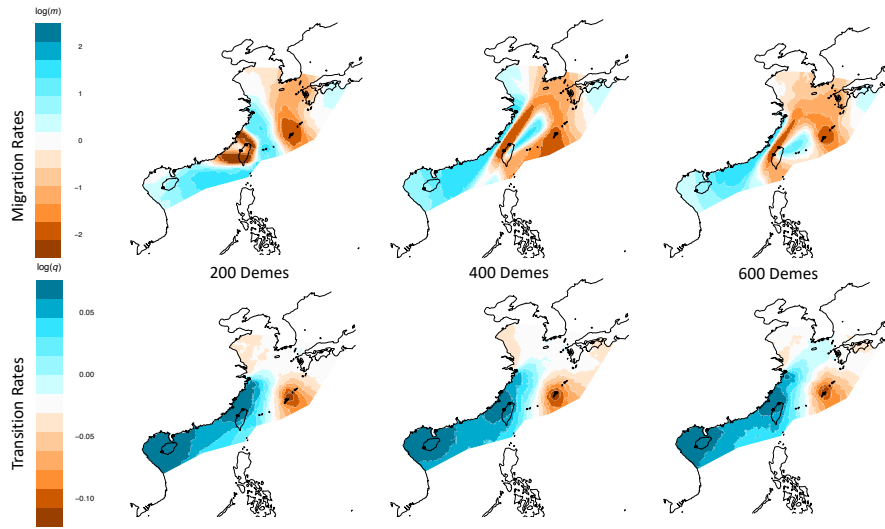


Figure 3.3 Migration and dissimilarity heatmaps.
Surface maps showing m , migration(top) and q , the genetic dissimilarity between two individuals within a deme (bottom). Simulations were run using 200,400 and 600 demes.

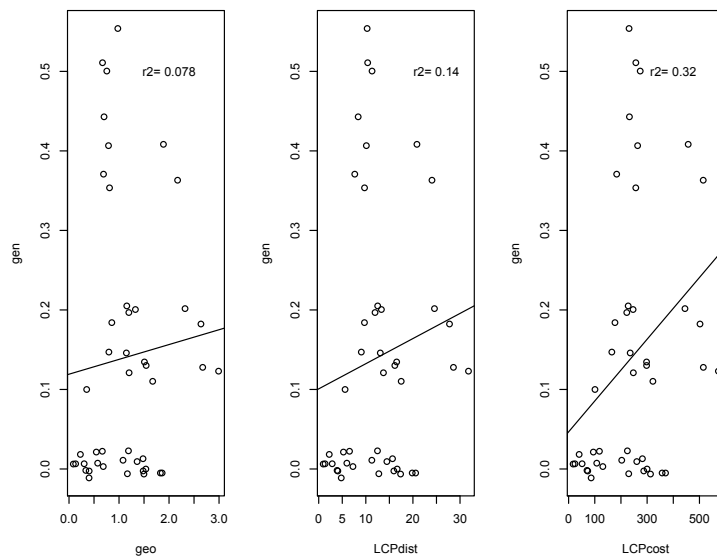


Figure 3.4 Genetic variation plotted against the three forms of geographic distance.

Table 3.1
Overall and population specific Heterozygosity (H_e)

Heterozygosity:		Overall: 0.0343		
Populations:				
Ariake	Seto	Okinawa	Taoyuan	Taichung
0.046	0.035	0.010	0.034	0.036
Cixi	Hainan	Xiapu	Xiashan	Zhoushan
0.033	0.037	0.036	0.034	0.037

Table 3.2 Pairwise F_{st} table.

	Ariake	Cixi	Hainan	Okinawa	Seto	Taoyuan	WestTaiwan	Xiapu	Xiashan	Zhoushan
Ariake	-	0.16	0.15	0.27	0.09	0.16	0.17	0.17	0.17	0.13
Cixi	0.16	-	-0.01	0.33	0.11	0.01	0.00	-0.01	-0.01	0.01
Hainan	0.15	-0.01	-	0.27	0.11	0.01	0.01	0.00	0.00	-0.01
Okinawa	0.27	0.33	0.27	-	0.36	0.31	0.26	0.29	0.29	0.34
Seto	0.09	0.11	0.11	0.36	-	0.12	0.10	0.11	0.12	0.13
Taoyuan	0.16	0.01	0.01	0.31	0.12	-	0.01	0.02	0.02	0.02
WestTaiwan	0.17	0.00	0.01	0.26	0.10	0.01	-	0.01	0.01	0.02
Xiapu	0.17	-0.01	0.00	0.29	0.11	0.02	0.01	-	-0.01	0.00
Xiashan	0.17	-0.01	0.00	0.29	0.12	0.02	0.01	-0.01	-	0.00
Zhoushan	0.13	0.01	-0.01	0.34	0.13	0.02	0.02	0.00	0.00	-

Pairwise F_{st} between all populations. Populations include two from Japan (Ariake and Seto), two from Taiwan (WestTaiwan and Taoyuan), Okinawa, and five from China (Cixi, Hainan, Xiapu, Xiashan, Zhoushan)

Table 3.3 Mantel and partial mantel test of genetic and geographic distances.

Mantel/Partial mantel test	Mantel r	P-value
F _{st} , EUC	0.07921	0.291
F _{st} , LCPD	0.1515	0.195
F _{st} , LCP	0.333	0.029
F _{st} , EUC, LCPD	-0.5847	1
F _{st} , EUC, LCP	-0.8323	1
F _{st} , LCPD, EUC	0.594	0.006
F _{st} , LCPD, LCP	-0.855	1
F _{st} , LCP, EUC	0.8515	0.001
F _{st} , LCP, LCPD	0.8691	0.001

Mantel and partial mantel test of genetic diversity ($F_{st}/1 - F_{st}$), Euclidian distance (EUC), and least cost path distance (LCPD) as well as a least cost path (LDP).

Table 3.4 ABC model selection

A) Proportion of accepted simulations (rejection):				
VC	VCe	SST	SSJ	IM
0.0722	0.8326	0.0458	0.0431	0.0063

B) Bayes factors				
	VC	SST	SSJ	IM
VCe	18.9500	26.1379	26.1379	189.5

A) The proportion of simulated models that produced summary statistic similar to the observed data for each scenario. B) likelihood comparison between model with most accepted simulation (VCe) and all other models. Models included an equilibrium island model (IM), stepping stone model with Okinawa being colonized from Taiwan (SST) or Okinawa being colonized from Japan (SSJ), vicariance/colonization models (VC), and vicariance/colonization models with exponential expansion in Okinawa (VCe).

Table 3.5 ABC parameter estimates

Parameter	Estimate	95% HPD	
N_C	73224	29945	99081
N_J	1675	199	6811
N_T	5666	786	9625
N_O	5372	400	9722
M_{CJ}	1.87E-04	6.77E-06	4.64E-04
M_{CT}	2.20E-04	1.01E-05	4.85E-04
M_{CO}	4.77E-05	2.64E-06	9.61E-05
M_{JC}	2.71E-04	1.46E-05	4.90E-04
M_{JT}	2.84E-04	1.51E-05	4.88E-04
M_{JO}	5.45E-05	3.42E-06	9.87E-05
M_{TC}	2.48E-04	8.66E-06	4.82E-04
M_{TJ}	1.93E-04	4.18E-06	4.77E-04
M_{TO}	5.36E-05	6.23E-06	9.69E-05
M_{OC}	5.67E-05	4.28E-06	9.85E-05
M_{OJ}	5.77E-05	2.88E-06	9.82E-05
M_{OT}	5.58E-05	3.72E-06	9.83E-05
T_{TC}	2890	1032	4667
T_{JC}	2924	1007	4824
T_{OC}	2857	1024	4863

Estimation of population size (N_i), migration rate from population i to j (M_{ij}) and time of coalescence in generations at which population i colonized form j (T_{ij}). Population include China (C), Japan (J), Okinawa (O), and Taiwan (T).

Appendix B – Additional Files

Sample Coverage

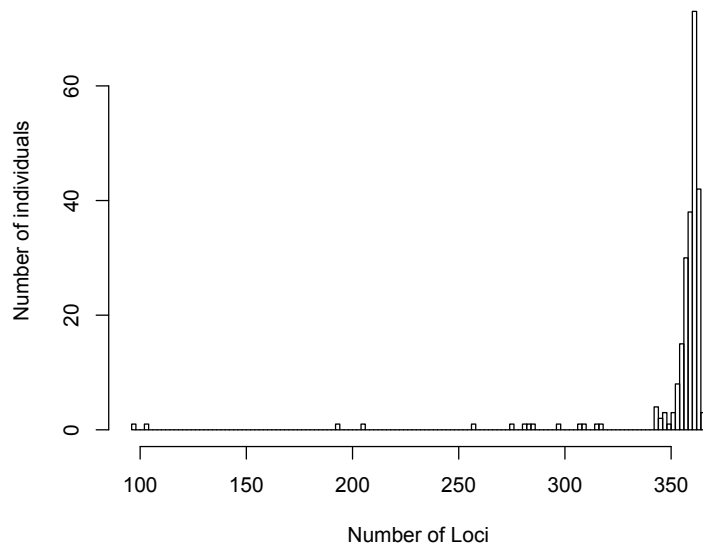


Figure A3.1
Histogram of frequency of number of loci per individual sample.

File A3.1

3RAD parameter file

```
----- ipyrad params file (v.0.5.15)-----
Master      ## [0] [assembly_name]: Assembly name. Used to name output directories for assembly steps
./          ## [1] [project_dir]: Project dir (made in curdir if not present)
           ## [2] [raw_fastq_path]: Location of raw non-demultiplexed fastq files
           ## [3] [barcodes_path]: Location of barcodes file
./postStep1/*fastq.gz ## [4] [sorted_fastq_path]: Location of demultiplexed/sorted fastq files
denovo      ## [5] [assembly_method]: Assembly method (denovo, reference, denovo+reference, denovo-reference)
           ## [6] [reference_sequence]: Location of reference sequence file
pairedrad   ## [7] [datatype]: Datatype (see docs): rad, gbs, ddrad, etc.
ATCGG,CGATCC ## [8] [restriction_overhang]: Restriction overhang (cut1,) or (cut1, cut2)
5           ## [9] [max_low_qual_bases]: Max low quality base calls (Q<20) in a read
33          ## [10] [phred_Qscore_offset]: phred Q score offset (33 is default and very standard)
6           ## [11] [mindepth_statistical]: Min depth for statistical base calling
6           ## [12] [mindepth_majrule]: Min depth for majority-rule base calling
10000       ## [13] [maxdepth]: Max cluster depth within samples
0.85        ## [14] [clust_threshold]: Clustering threshold for de novo assembly
1           ## [15] [max_barcode_mismatch]: Max number of allowable mismatches in barcodes
2           ## [16] [filter_adapters]: Filter for adapters/primers (1 or 2=striker)
35          ## [17] [filter_min_trim_len]: Min length of reads after adapter trim
2           ## [18] [max_alleles_consens]: Max alleles per site in consensus sequences
5, 5        ## [19] [max_Ns_consens]: Max N's (uncalled bases) in consensus (R1, R2)
8, 8        ## [20] [max_Hs_consens]: Max Hs (heterozygotes) in consensus (R1, R2)
12          ## [21] [min_samples_locus]: Min # samples per locus for output
5, 5        ## [22] [max_SNPs_locus]: Max # SNPs per locus (R1, R2)
8, 8        ## [23] [max_Indels_locus]: Max # of indels per locus (R1, R2)
0.5         ## [24] [max_shared_Hs_locus]: Max # heterozygous sites per locus (R1, R2)
0, 0        ## [25] [edit_cutsites]: Edit cut-sites (R1, R2) (see docs)
0, 0, 0, 0  ## [26] [trim_overhang]: Trim overhang (see docs) (R1>, <R1, R2>, <R2)
l, p, s, v  ## [27] [output_formats]: Output formats (see docs)
           ## [28] [pop_assign_file]: Path to population assignment file
params-Master_2.txt (END)
```

File A3.2

RADcap parameter file

```
----- ipyrad params file (v.0.7.21)-----
restr      ## [0] [assembly_name]: Assembly name. Used to name output directories for assembly steps
/Volumes/JoelCorush5/2018_Jan_HiSeq/for_Joel      ## [1] [project_dir]: Project dir (made in curdir if not present)
      ## [2] [raw_fastq_path]: Location of raw non-demultiplexed fastq files
      ## [3] [barcodes_path]: Location of barcodes file
/Volumes/JoelCorush5/2018_Jan_HiSeq/for_Joel/decloned/*R*      ## [4] [sorted_fastq_path]: Location of demultiplexed/sorted fastq
files
reference   ## [5] [assembly_method]: Assembly method (denovo, reference, denovo+reference, denovo-reference)
/Volumes/JoelCorush5/2018_Jan_HiSeq/for_Joel/ref/PE_ref.fasta ## [6] [reference_sequence]: Location of reference sequence file
pairedrad   ## [7] [datatype]: Datatype (see docs): rad, gbs, ddrad, etc.
CGG,GATCC   ## [8] [restriction_overhang]: Restriction overhang (cut1,) or (cut1, cut2)
5           ## [9] [max_low_qual_bases]: Max low quality base calls (Q<20) in a read
33          ## [10] [phred_Qscore_offset]: phred Q score offset (33 is default and very standard)
6           ## [11] [mindepth_statistical]: Min depth for statistical base calling
6           ## [12] [mindepth_majrule]: Min depth for majority-rule base calling
10000       ## [13] [maxdepth]: Max cluster depth within samples
0.85        ## [14] [clust_threshold]: Clustering threshold for de novo assembly
0           ## [15] [max_barcode_mismatch]: Max number of allowable mismatches in barcodes
1           ## [16] [filter_adapters]: Filter for adapters/primers (1 or 2=stricter)
35          ## [17] [filter_min_trim_len]: Min length of reads after adapter trim
2           ## [18] [max_alleles_consens]: Max alleles per site in consensus sequences
5, 5        ## [19] [max_Ns_consens]: Max N's (uncalled bases) in consensus (R1, R2)
8, 8        ## [20] [max_Hs_consens]: Max Hs (heterozygotes) in consensus (R1, R2)
220         ## [21] [min_samples_locus]: Min # samples per locus for output
50, 50      ## [22] [max_SNPs_locus]: Max # SNPs per locus (R1, R2)
15, 15      ## [23] [max_Indels_locus]: Max # of indels per locus (R1, R2)
0.5         ## [24] [max_shared_Hs_locus]: Max # heterozygous sites per locus (R1, R2)
0, 0, 0, 0  ## [25] [trim_reads]: Trim raw read edges (R1>, <R1, R2>, <R2) (see docs)
0, 0, 0, 0  ## [26] [trim_loci]: Trim locus edges (see docs) (R1>, <R1, R2>, <R2)
G, a, g, k, m, l, n, p, s, u, t, v      ## [27] [output_formats]: Output formats (see docs)
      ## [28] [pop_assign_file]: Path to population assignment file
```

File A3.3

strataG_fsc_2ec.R file – Vicariance/colonization with exponential expansion of Okinawa

```
## China, Japan, Taiwan, Okinawa
# Okinawa(3), Taiwan(2), and Japan(1) arise independently from China(0)
#
## parameters
# population sizes
# migration rates
# mutation rates
# colonization event times

reps <- 1000

inparams <- matrix(nrow=reps,ncol=4+12+3)
colnames(inparams) <-
c("N1","N2","N3","N4","m12","m13","m14","m21","m23","m24","m31","m32","m34","m41","m42","m43","t20","t10","t30")
sumstats <- matrix(nrow=reps,ncol=15)
colnames(sumstats) <- c("Fst","f1","f2","f3","f4","f5","f6","l1","l2","l3","l4","h1","h2","h3","h4")

for(i in 1:reps){
  pop.size <- round(c(runif(1,1e+2,1e+5),runif(1,1e+2,1e+4),runif(1,1e+2,1e+4),runif(1,1e+2,1e+4)))

  #exponential expansion of Okinawa
  num.gen <- round(runif(3, 900, 5000))
  growth.rate <- c(0,0,0,log(100/pop.size[4])/num.gen[3])

  pop.info <- fscPopInfo(pop.size=pop.size, sample.size=c(119,55,50,12), growth.rate=growth.rate)

  # sample migration rates from uniform distributions loosely based on Fst=1/(Nm + 1)
  mig.rates <- matrix(c(0,runif(1,0,5e-4), runif(1,0,5e-4), runif(1,0,1e-4),
    runif(1,0,5e-4), 0, runif(1,0,5e-4), runif(1,0,1e-4),
    runif(1,0,5e-4), runif(1,0,5e-4),0, runif(1,0,1e-4),
    runif(1,0,1e-4), runif(1,0,1e-4), runif(1,0,1e-4), 0),ncol=4)

  # mutation rate of 1e-7 yields about 365 variable loci of 1300 simulated SNPs
  snp.params <- fscLocusParams(locus.type="snp", num.loci=1, mut.rate=rep(1e-7,1300)) # sample mutation rates from a prior?

  # historical events
  #2: Taiwan(2) arises from China(0), Japan(1) arises from China(0), Okinawa(3) arises from China(0)
  source.deme <- c(2,1,3)
  sink.deme <- c(0,0,0)
  prop.migrants <- 1
  new.sink.size <- 1
  new.sink.growth <- 0
  new.mig.mat <- 0
  hist.ev <- cbind(num.gen = num.gen, source.deme = source.deme, sink.deme = sink.deme,
    prop.migrants = prop.migrants, new.sink.size = new.sink.size,
    new.sink.growth = new.sink.growth, new.mig.mat = new.mig.mat)

  # record inputs for ABC
  inparams[i,] <- c(pop.size,as.vector(mig.rates)[-c(1,6,11,16)],num.gen)

  # run fastsimcoal
  sim.snps <- fastsimcoal(pop.info, snp.params, mig.rates, hist.ev=hist.ev, exec="fsc26", delete.files=TRUE)

  genin <- gtypes2genind(sim.snps) # simulated data in genind format

  # drop invariant loci (only record if greater than 2)
  x <- genin$loc.n.all > 1
  if(sum(x)>2){
    genin <- genin[loc=x]
    hfstat <- genind2hierfst(genin) # simulated data in hierfst format
  }
```

```

# summary statistics
AR <- colMeans(allelic.richness(hfstat)$Ar,na.rm=TRUE) # allelic richness by population
het <- Hs(genin) # average gene diversity by population
Fst <- wc(hfstat)$FST # global Fst
pairw <- as.vector(as.dist(pairwise.wcFst(hfstat))) # pairwise Fst

sumstats[i,] <- c(Fst,pairw,AR,het)
}

# save simulated inputs and summary statistics
write.table(inparams, file="inparams2ec.txt",append=TRUE, col.names=FALSE)
write.table(sumstats, file="sumstats2ec.txt",append=TRUE, col.names=FALSE)

```

File A3.4

strataG_fsc_1eq.R file – Equilibrium island model

```
library("strataG")
library("hierfstat")
library("pegas")
library(abc)

## China, Japan, Taiwan, Okinawa
# equilibrium island model
sumstat <- c(0.12192541, 0.04633412, 0.14038424, 0.23861298, 0.12298461, 0.22605168, 0.27133118, 1.03317441, 1.03443200,
1.02725645, 1.01277428, 0.03161648, 0.03937682, 0.03233235, 0.01764387)
#
## parameters
# population sizes
# migration rates
# mutation rates
# colonization event times

reps <- 1000

inparams <- matrix(nrow=reps,ncol=4+12)
colnames(inparams) <- c("N1","N2","N3","N4","m12","m13","m14","m21","m23","m24","m31","m32","m34","m41","m42","m43")
sumstats <- matrix(nrow=reps,ncol=15)
colnames(sumstats) <- c("Fst","f1","f2","f3","f4","f5","f6","I1","I2","I3","I4","h1","h2","h3","h4")

for(i in 1:reps){
  # sample population sizes from uniform distributions loosely based on preliminary runs
  pop.size <- round(c(runif(1,1e+2,1e+4),runif(1,1e+2,1e+4),runif(1,1e+2,1e+4),runif(1,1e+2,1e+4)))

  # exponential expansion of island pops
  # T1 <- round(runif(1,100,5000))
  # T2 <- round(runif(1,T1,7000))
  # T3 <- round(runif(1,100,7000))
  # num.gen <- c(T1,T2,T3)
  # growth.rate <- c(0,log(100/pop.size[-1])/num.gen)
  growth.rate <- 0

  pop.info <- fscPopInfo(pop.size=pop.size, sample.size=c(119,55,50,12), growth.rate=growth.rate)

  # sample migration rates from uniform distributions loosely based on  $F_{st}=1/(Nm + 1)$ 
  mig.rates <- matrix(c(0,runif(1,0,5e-4), runif(1,0,5e-4), runif(1,0,1e-4),
runif(1,0,5e-4), 0, runif(1,0,5e-4), runif(1,0,1e-4),
runif(1,0,5e-4), runif(1,0,5e-4),0, runif(1,0,1e-4),
runif(1,0,1e-4), runif(1,0,1e-4), runif(1,0,1e-4), 0),ncol=4)

  # mutation rate of 1e-7 yields about 365 variable loci of 1300 simulated SNPs
  snp.params <- fscLocusParams(locus.type="snp", num.loci=1, mut.rate=rep(1e-7,1300)) # sample mutation rates from a
prior?

  # record inputs for ABC
  inparams[i,] <- c(pop.size,as.vector(mig.rates)[-c(1,6,11,16)])

  # run fastsimcoal
  sim.snps <- fastsimcoal(pop.info, snp.params, mig.rates, hist.ev=NULL, exec="fsc26", delete.files=TRUE)

  genin <- gtypes2genind(sim.snps) # simulated data in genind format

  # drop invariant loci (only record if greater than 2)
  x <- genin$loc.n.all > 1
  if(sum(x)>2){
    genin <- genin[loc=x]
    hfstat <- genind2hierfst(genin) # simulated data in hierfstat format
  }
```

```

# summary statistics
AR <- colMeans(allelic.richness(hfstat)$Ar,na.rm=TRUE) # allelic richness by population
het <- Hs(genin) # average gene diversity by population
Fst <- wc(hfstat)$FST # global Fst
pairw <- as.vector(as.dist(pairwise.wcFst(hfstat))) # pairwise Fst

sumstats[i,] <- c(Fst,pairw,AR,het)}
}

# save simulated inputs and summary statistics
write.table(inparams, file="inparams1eq.txt",append=TRUE, col.names=FALSE)
write.table(sumstats, file="sumstats1eq.txt",append=TRUE, col.names=FALSE)

```

File A3.5

strataG_fsc_2e.R file – Vicariance/colonization

```
library("strataG")
library("hierfstat")
library("pegas")
library(abc)

## China, Japan, Taiwan, Okinawa
# Okinawa(3), Taiwan(2), and Japan(1) arise independently from China(0)
sumstat <- c(0.12192541, 0.04633412, 0.14038424, 0.23861298, 0.12298461, 0.22605168, 0.27133118, 1.03317441, 1.03443200,
1.02725645, 1.01277428, 0.03161648, 0.03937682, 0.03233235, 0.01764387)
#
## parameters
# population sizes
# migration rates
# mutation rates
# colonization event times

reps <- 1000

inparams <- matrix(nrow=reps,ncol=4+12+3)
colnames(inparams) <-
c("N1","N2","N3","N4","m12","m13","m14","m21","m23","m24","m31","m32","m34","m41","m42","m43","t20","t10","t30")
sumstats <- matrix(nrow=reps,ncol=15)
colnames(sumstats) <- c("Fst","f1","f2","f3","f4","f5","f6","l1","l2","l3","l4","h1","h2","h3","h4")

for(i in 1:reps){
  # sample population sizes from uniform distributions loosely based on preliminary runs
  pop.size <- round(c(runif(1,1e+2,1e+4),runif(1,1e+2,1e+4),runif(1,1e+2,1e+4),runif(1,1e+2,1e+4)))

  #exponential expansion of island pops
  num.gen <- round(runif(3, 900, 5000))
  growth.rate <- c(0,log(100/pop.size[-1])/num.gen)

  pop.info <- fscPopInfo(pop.size=pop.size, sample.size=c(119,55,50,12), growth.rate=growth.rate)

  # sample migration rates from uniform distributions loosely based on Fst=1/(Nm + 1)
  mig.rates <- matrix(c(0,runif(1,0,5e-4), runif(1,0,5e-4), runif(1,0,1e-4),
runif(1,0,5e-4), 0, runif(1,0,5e-4), runif(1,0,1e-4),
runif(1,0,5e-4), runif(1,0,5e-4),0, runif(1,0,1e-4),
runif(1,0,1e-4), runif(1,0,1e-4), runif(1,0,1e-4), 0),ncol=4)

  # mutation rate of 1e-7 yields about 365 variable loci of 1300 simulated SNPs
  snp.params <- fscLocusParams(locus.type="snp", num.loci=1, mut.rate=rep(1e-7,1300)) # sample mutation rates from a
prior?

  # historical events
  #2: Taiwan(2) arises from China(0), Japan(1) arises from China(0), Okinawa(3) arises from China(0)
  source.deme <- c(2,1,3)
  sink.deme <- c(0,0,0)
  prop.migrants <- 1
  new.sink.size <- 1
  new.sink.growth <- 0
  new.mig.mat <- 0
  hist.ev <- cbind(num.gen = num.gen, source.deme = source.deme, sink.deme = sink.deme,
prop.migrants = prop.migrants, new.sink.size = new.sink.size,
new.sink.growth = new.sink.growth, new.mig.mat = new.mig.mat)

  # record inputs for ABC
  inparams[i,] <- c(pop.size,as.vector(mig.rates)[-c(1,6,11,16)],num.gen)

  # run fastsimcoal
```

```

sim.snps <- fastsimcoal(pop.info, snp.params, mig.rates, hist.ev=hist.ev, exec="fsc26", delete.files=TRUE)

genin <- gtypes2genind(sim.snps) # simulated data in genind format

# drop invariant loci (only record if greater than 2)
x <- genin$loc.n.all > 1
if(sum(x)>2){
  genin <- genin[loc=x]
  hfstat <- genind2hierfst(genin) # simulated data in hierfst format

# summary statistics
AR <- colMeans(allelic.richness(hfstat)$Ar,na.rm=TRUE) # allelic richness by population
het <- Hs(genin) # average gene diversity by population
Fst <- wc(hfstat)$FST # global Fst
  pairw <- as.vector(as.dist(pairwise.wcFst(hfstat))) # pairwise Fst

  sumstats[i,] <- c(Fst,pairw,AR,het)}
}

# save simulated inputs and summary statistics
write.table(inparams, file="inparams2e.txt",append=TRUE, col.names=FALSE)
write.table(sumstats, file="sumstats2e.txt",append=TRUE, col.names=FALSE)

```

File A3.6

strataG_fsc_3e.R file – Stepping stone – Okinawa colonized from Taiwan

```
library("strataG")
library("hierfstat")
library("pegas")
library(abc)

## China, Japan, Taiwan, Okinawa
# Okinawa(3) arises from Taiwan(2), Taiwan(2) arises from China(0), Japan(1) arises from China(0)
sumstat <- c(0.12192541, 0.04633412, 0.14038424, 0.23861298, 0.12298461, 0.22605168, 0.27133118, 1.03317441, 1.03443200,
1.02725645, 1.01277428, 0.03161648, 0.03937682, 0.03233235, 0.01764387)
#
## parameters
# population sizes
# migration rates
# mutation rates
# colonization event times

reps <- 1000

inparams <- matrix(nrow=reps,ncol=4+12+3)
colnames(inparams) <-
c("N1","N2","N3","N4","m12","m13","m14","m21","m23","m24","m31","m32","m34","m41","m42","m43","t20","t10","t30")
sumstats <- matrix(nrow=reps,ncol=15)
colnames(sumstats) <- c("Fst","f1","f2","f3","f4","f5","f6","l1","l2","l3","l4","h1","h2","h3","h4")

for(i in 1:reps){
  # sample population sizes from uniform distributions loosely based on preliminary runs
  pop.size <- round(c(runif(1,1e+2,1e+4),runif(1,1e+2,1e+4),runif(1,1e+2,1e+4),runif(1,1e+2,1e+4)))

  #exponential expansion of island pops
  T1 <- round(runif(1,100,5000))
  T2 <- round(runif(1,T1,7000))
  T3 <- round(runif(1,100,7000))
  num.gen <- c(T1,T2,T3)
  growth.rate <- c(0,log(100/pop.size[-1])/num.gen)

  pop.info <- fscPopInfo(pop.size=pop.size, sample.size=c(119,55,50,12), growth.rate=growth.rate)

  # sample migration rates from uniform distributions loosely based on  $F_{st}=1/(N_m + 1)$ 
  mig.rates <- matrix(c(0,runif(1,0,5e-4), runif(1,0,5e-4), runif(1,0,1e-4),
runif(1,0,5e-4), 0, runif(1,0,5e-4), runif(1,0,1e-4),
runif(1,0,5e-4), runif(1,0,5e-4),0, runif(1,0,1e-4),
runif(1,0,1e-4), runif(1,0,1e-4), runif(1,0,1e-4), 0),ncol=4)

  # mutation rate of 1e-7 yields about 365 variable loci of 1300 simulated SNPs
  snp.params <- fscLocusParams(locus.type="snp", num.loci=1, mut.rate=rep(1e-7,1300)) # sample mutation rates from a
prior?

  # historical events
# Okinawa(3) arises from Taiwan(2), Taiwan(2) arises from China(0), Japan(1) arises from China(0)
  source.deme <- c(3,2,1)
  sink.deme <- c(2,0,0)
  prop.migrants <- 1
  new.sink.size <- 1
  new.sink.growth <- 0
  new.mig.mat <- 0
  hist.ev <- cbind(num.gen = num.gen, source.deme = source.deme, sink.deme = sink.deme,
  prop.migrants = prop.migrants, new.sink.size = new.sink.size,
  new.sink.growth = new.sink.growth, new.mig.mat = new.mig.mat)
```

record inputs for ABC

```

inparams[i,] <- c(pop.size,as.vector(mig.rates)[-c(1,6,11,16)],num.gen)

# run fastsimcoal
sim.snps <- fastsimcoal(pop.info, snp.params, mig.rates, hist.ev=hist.ev, exec="fsc26", delete.files=TRUE)

genin <- gtypes2genind(sim.snps) # simulated data in genind format

# drop invariant loci (only record if greater than 2)
x <- genin$loc.n.all > 1
if(sum(x)>2){
  genin <- genin[loc=x]
  hfstat <- genind2hierfst(genin) # simulated data in hierfst format

# summary statistics
AR <- colMeans(allelic.richness(hfstat)$Ar,na.rm=TRUE) # allelic richness by population
het <- Hs(genin) # average gene diversity by population
Fst <- wc(hfstat)$FST # global Fst
  pairw <- as.vector(as.dist(pairwise.wcFst(hfstat))) # pairwise Fst

  sumstats[i,] <- c(Fst,pairw,AR,het)}
}

# save simulated inputs and summary statistics
write.table(inparams, file="inparams3e.txt",append=TRUE, col.names=FALSE)
write.table(sumstats, file="sumstats3e.txt",append=TRUE, col.names=FALSE)

```

File A3.7

strataG_fsc_4e.R file – Stepping stone – Okinawa colonized from Japan

```
library("strataG")
library("hierfstat")
library("pegas")
library(abc)

## China, Japan, Taiwan, Okinawa
# Okinawa(3) arises from Japan(1), Taiwan(2) arises from China(0), Japan(1) arises from China(0)
sumstat <- c(0.12192541, 0.04633412, 0.14038424, 0.23861298, 0.12298461, 0.22605168, 0.27133118, 1.03317441, 1.03443200,
1.02725645, 1.01277428, 0.03161648, 0.03937682, 0.03233235, 0.01764387)
#
## parameters
# population sizes
# migration rates
# mutation rates
# colonization event times

reps <- 1000

inparams <- matrix(nrow=reps, ncol=4+12+3)
colnames(inparams) <-
c("N1", "N2", "N3", "N4", "m12", "m13", "m14", "m21", "m23", "m24", "m31", "m32", "m34", "m41", "m42", "m43", "t20", "t10", "t30")
sumstats <- matrix(nrow=reps, ncol=15)
colnames(sumstats) <- c("Fst", "f1", "f2", "f3", "f4", "f5", "f6", "l1", "l2", "l3", "l4", "h1", "h2", "h3", "h4")

for(i in 1:reps){
  # sample population sizes from uniform distributions loosely based on preliminary runs
  pop.size <- round(c(runif(1, 1e+2, 1e+4), runif(1, 1e+2, 1e+4), runif(1, 1e+2, 1e+4), runif(1, 1e+2, 1e+4)))

  #exponential expansion of island pops
  T1 <- round(runif(1, 100, 5000))
  T2 <- round(runif(1, T1, 7000))
  T3 <- round(runif(1, 100, 7000))
  num.gen <- c(T1, T2, T3)
  growth.rate <- c(0, log(100/pop.size[-1])/num.gen)

  pop.info <- fscPopInfo(pop.size=pop.size, sample.size=c(119, 55, 50, 12), growth.rate=growth.rate)

  # sample migration rates from uniform distributions loosely based on Fst=1/(Nm + 1)
  mig.rates <- matrix(c(0, runif(1, 0, 5e-4), runif(1, 0, 5e-4), runif(1, 0, 1e-4),
    runif(1, 0, 5e-4), 0, runif(1, 0, 5e-4), runif(1, 0, 1e-4),
    runif(1, 0, 5e-4), runif(1, 0, 5e-4), 0, runif(1, 0, 1e-4),
    runif(1, 0, 1e-4), runif(1, 0, 1e-4), runif(1, 0, 1e-4), 0), ncol=4)

  # mutation rate of 1e-7 yields about 365 variable loci of 1300 simulated SNPs
  snp.params <- fscLocusParams(locus.type="snp", num.loci=1, mut.rate=rep(1e-7, 1300)) # sample mutation rates from a
prior?

  # historical events
# Okinawa(3) arises from Japan(1), Taiwan(2) arises from China(0), Japan(1) arises from China(0)
  source.deme <- c(3, 1, 2)
  sink.deme <- c(1, 0, 0)
  prop.migrants <- 1
  new.sink.size <- 1
  new.sink.growth <- 0
  new.mig.mat <- 0
  hist.ev <- cbind(num.gen = num.gen, source.deme = source.deme, sink.deme = sink.deme,
    prop.migrants = prop.migrants, new.sink.size = new.sink.size,
    new.sink.growth = new.sink.growth, new.mig.mat = new.mig.mat)

  # record inputs for ABC
  inparams[i,] <- c(pop.size, as.vector(mig.rates)[-c(1, 6, 11, 16)], num.gen)
```

```

# run fastsimcoal
sim.snps <- fastsimcoal(pop.info, snp.params, mig.rates, hist.ev=hist.ev, exec="fsc26", delete.files=TRUE)

genin <- gtypes2genind(sim.snps) # simulated data in genind format

# drop invariant loci (only record if greater than 2)
x <- genin$loc.n.all > 1
if(sum(x)>2){
  genin <- genin[loc=x]
  hfstat <- genind2hierfst(genin) # simulated data in hierfst format

  # summary statistics
  AR <- colMeans(allelic.richness(hfstat)$Ar,na.rm=TRUE) # allelic richness by population
  het <- Hs(genin) # average gene diversity by population
  Fst <- wc(hfstat)$FST # global Fst
  pairw <- as.vector(as.dist(pairwise.wcFst(hfstat))) # pairwise Fst

  sumstats[i,] <- c(Fst,pairw,AR,het)}
}

# save simulated inputs and summary statistics
write.table(inparams, file="inparams4e.txt",append=TRUE, col.names=FALSE)
write.table(sumstats, file="sumstats4e.txt",append=TRUE, col.names=FALSE)

```

CONCLUSION

Evolutionary patterns of alternative life histories in fishes seems to be a complex subject that requires a wide array of questions, methods and viewpoints to be addressed. The patterns, mechanisms, and causes of each independent gain, loss, and persistence of these life histories each warrant their own investigations. This research has contributed to ongoing understanding of these life histories by 1) identifying patterns based on previous hypotheses about transitions in and out of diadromy, 2) assessing a potential mechanism of the pattern observed, and 3) looking at a second life history, amphibious behavior, to identify a mechanism that causes the life history to affect population structure. These three independent, yet compatible, aspects of life history evolution help to demonstrate the complexity of individual behavioral shifts and the effects of those behaviors on patterns of biodiversity on past and current biodiversity.

VITA

Joel B. Corush was born in Skokie Illinois in 1985. He attended Niles North High School. He began his research career at Drake University looking at the ecology of frog nesting behavior with Dr. James Learned Christiansen as well as variation in stream fish communities in central Iowa and the effects of nitrate on gill morphology of Fathead Minnows, *Pimephales promelas* with Dr. Thomas Rosburg. He was awarded his bachelor's degree in Biochemistry Cellular and Molecular Biology from Drake University in 2008. Joel then worked at the Field Museum where he assisted with the Botany collection under the supervision of Dr. Matt von Konrat and studied the molecular evolution of Lichens with Dr. Thorsten Lumbsch. He then worked in multiple labs including the lab of Dr. John Epifanio at The Illinois Natural History Survey to study hybridization in freshwater sunfish, the lab of Dr. Nelson Ting at University of Iowa investigation population genetic of old world monkeys, the lab of Dr. Mark Westneat at the Field Museum conducting molecular lab work and finally the lab of Dr. David Lodge at Notre Dame where he helped develop methods for environmental DNA (eDNA) detection in Asian Carp in the great lakes. In the fall of 2012, Joel joined the lab of Dr. Benjamin Fitzpatrick at The University of Tennessee for his Ph.D.