2017

Go with the flow: patterns of connectivity in low dispersal coral reef gobies (Coryphopterus spp.) throughout the western Atlantic

Daniel Volk
University of Central Florida

GO WITH THE FLOW: PATTERNS OF CONNECTIVITY IN LOW DISPERSAL CORAL REEF GOBIES (CORYPHOPTERUS SPP.) THROUGHOUT THE WESTERN ATLANTIC

by
DANIEL R. VOLK
B.S. Cleveland State University, 2015

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Fall Term
2017

Major Professor: Eric A. Hoffman

ABSTRACT

Gene flow is an integral biological process that can mediate speciation. While many consider the ocean to be an open environment, there are many barriers that limit gene flow, particularly in the western Atlantic. I analyzed data from two widespread, coral reef fishes, the bridled goby (*Coryphopterus glaucofraenum*) and sand-canyon goby (*C. venezuelae*), throughout their range in the western Atlantic. Using two genetic datasets, mitochondrial DNA (mtDNA) and genomic SNPs, I investigated the evolutionary history of these species and inferred the location and strength of putative barriers. My results suggest that several unique lineages have genetically diverged from one another in the presence of two major barriers. First, the Amazon River has isolated Brazil from the Caribbean and second, a unique lineage was found at an isolated oceanic island, Atol das Rocas, off the northeast coast of Brazil. Furthermore, minor barriers have caused slight genetic differentiation in each of the Caribbean species off the coast of Venezuela, while on the Brazilian coast, there are up to two barriers that separate three genetically unique areas. The stronger of the two barriers is located at Cabo Frio near an upwelling system and the weaker barrier coincides with the outflow of the São Francisco River. Overall, this research highlights how barriers impact speciation and genetic structure within these gobies in the western Atlantic and more broadly, deepens our understanding about the role of oceanographic features in the speciation process.

ACKNOWLEDGMENTS

I have grown immensely as a researcher and immeasurably as a person since arriving in Orlando. I would first like to thank my advisor, Eric Hoffman. With your patience and caring nature, you fostered an incredible environment for learning and I am sincerely grateful for that. You also literally took me halfway around the world and introduced me to countless foods that I would not have experienced without your recommendation. Geoff Cook, you were quick to provide assistance with areas that I am less than comfortable with, such as statistics and oceanography. In fact, I think we had more full-length conversations in the hallway than in a sit-down manner which shows your easygoing personality. I also want to thank Cadu for welcoming me as an honorary member of your lab and introducing me to the world of marine phylogeography. I am appreciative for Anna Savage and Barb Sharanowski who were kind enough to provide help with genomics and phylogenetics on multiple occasions. Carole Baldwin and Benjamin Victor provided valuable samples to me that I would not have obtained otherwise. The UCF Biology Department provided me with the opportunity to travel to amazing places, present my work, and meet so many wonderful people. I also want to thank my labmates, Lauren Caspers, Shelly Gaynor, Alexa Trujillo, Sávio Calazans, Kim Sheffield, Brittney Lebert, and Johnny Konvalina who helped me grow significantly as a person and provided valuable feedback on countless presentations and drafts. The Biology Graduate Student Association has been source of encouragement through life-long friendships that I hope to continue. Specifically, I sincerely want to recognize Ryan Chabot, Rhett Rautsaw, Havalend Steinmuller, Ian Kutch, Amber Jones,

Miles Zhang, Katie Mercier, Matthew Lawrance, Andrew Mason, Jason Strickland, Tiffani Manteuffel, Molly Grace and Katrina Phillips for inspiring your intense passion for your work and life into my own. To my Florida family, you have opened your doors to me many times and I cannot express my gratitude for your kind acts. My Ohio family has been waiting patiently for me to return home and that time has come. Lastly, I cannot put into words the support I have received from my fiancé Monica. You have sacrificed many things on my journey to obtain my graduate degree and I can only hope to repay you in a similar manner.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

Gene flow is an important biological process that can mediate genetic divergence of taxa such that continuous gene flow often results in homogenization, while isolation can lead to taxon divergence (Slatkin, 1985, 1987). It is typically easier to distinguish two species when gene flow is not occurring, however, it becomes difficult to differentiate species when intermittent gene flow occurs (Roux et al., 2016). While speciation represents an extreme result of isolation, reduced connectivity may lead to population-level differentiation as seen with reduced gene flow between subspecies of tigers over nearly one hundred thousand years (*Panthera tigris*; Luo et al., 2004). Similarly, recently isolated populations also exhibit patterns of population-level differentiation, such as have been found in Key deer (*Odocoileus virginianus clavium*) after the Florida Keys became separated from the mainland 6,000 – 10,000 years ago (Villanova, Hughes, & Hoffman, 2017).

Both isolated populations and species in which intermittent gene flow occurs often contain unique genetic characteristics that are difficult to detect using a small number of genetic markers (Spinks, Thomson, & Shaffer, 2014). These hidden genetic characteristics often lead to uncertainty when differentiating between taxa (Sukumaran & Knowles, 2017). Recently, genomic tools have been developed to detect subtle or recent divergences, such as the use of single-nucleotide polymorphisms (SNPs; Andrews, Good, Miller, Luikart, & Hohenlohe, 2016; Gaither et al., 2015; Gleason & Burton, 2016; Gottscho et al., 2017; Momigliano et al., 2017; Prates et al., 2016; Saenz-Agudelo et al., 2015). As a result of the increased power with genomic data, SNPs were able to distinguish population structure in

the endangered western pond turtle (*Emys marmorata*) when mtDNA and nuclear loci provided contrasting signals (Spinks, Thomson, & Shaffer, 2014).

Although marine populations often appear to be connected over vast distances due to the utilization of ocean currents for larval dispersal and the apparent absence of impermeable barriers (Palumbi, 1994), life-history characteristics can limit overall dispersal potential such that impediments become apparent in certain taxa. To illustrate how some taxa are more affected by barriers than others, pelagic spawning taxa disperse gametes through the water column for long-distance dispersal whereas demersal spawners lay their eggs directly on substrate which may be less influenced by currents (Blaxter, 2010; Gaylor and Gaines, 2000). These spawning modes may impact overall dispersal potential such that demersal spawning fishes can exhibit more structure among populations than pelagic spawning fishes (Bradbury, Laurel, Snelgrove, Bentzen, & Campana, 2008; Floeter et al., 2008; Riginos, Douglas, Jin, Shanahan, & Treml, 2011). Using ocean currents for dispersal means that gene flow primarily occurs in the direction of current; this trend is seen in several marine taxa including fishes (D'Agostini, Gherardi, & Pezzi, 2015; White et al., 2010), oysters (Faust et al., 2017) and hydrothermal vent tubeworms (Young, Fujio, & Vrijenhoek, 2008).

In contrast to ocean currents that facilitate gene flow, currents can diverge in two opposite directions and act as permeable barriers that prevent populations from interacting (Gaylord & Gaines, 2000). Often times, minor genetic isolation can occur in the presence of diverging currents due to larvae that are unable to disperse against flowing water (Yamazaki et al., 2017). For instance, populations of rabbitfish (*Siganus fuscescens*)

along the eastern Philippines exhibit genetic structure when an equatorial current splits into two opposite flowing currents (Magsino & Juinio-Meñez, 2008). Similarly, minor population isolation due to diverging currents has been found in king weakfish (*Macrodon ancylodon*) in the Atlantic (Santos, Hrbek, Farias, Schneider, & Sampaio, 2006).

Several currents throughout the western Atlantic impact connectivity among marine taxa such that common phylogeographic divisions occur throughout this region. Many empirical studies and oceanographic models have found that currents cause separation between the east and west Caribbean in addition to strong isolation of the Bahamas (Cowen, Paris, & Srinivasan, 2006; DeBiasse, Richards, Shivji, & Hellberg, 2016; Foster et al., 2012; Jackson et al., 2014; Taylor & Hellberg, 2003, 2006). Within Brazil, populations of fishes, crustaceans and corals often genetically group into three areas that roughly correspond to the southern equatorial current (SEC) bifurcation and the current-driven Cabo Frio upwelling (Boschi, 2000; Cunha, Souza, & Dias, 2014; Fernandes, Alves, Barros-alves, & Teixeira, 2012; Machado et al., 2017; Maggioni, Rogers, & Maclean, 2003; Picciani, de Lossio e Seiblitz, de Paiva, e Castro, & Zilberberg, 2016; Santos et al., 2006).

In addition to ocean currents that impact connectivity of marine species, freshwater and sediment outflow from rivers can reduce gene flow among taxa in the western Atlantic. The immense outflow from the Amazon River is carried north along the South American coast, which reduces salinity and increases sedimentation for thousands of kilometers (Ffield, 2007). In turn, this reduced salinity is known to cause speciation for many low dispersal marine taxa that are unable to traverse low salinity habitat (Floeter et al., 2008; Rocha, 2003). In fact, the Caribbean and Brazil are considered distinct biogeographic

provinces due to the isolation of each province caused by the Amazon barrier (Briggs & Bowen, 2012; Floeter et al., 2008).

Genetic isolation can also result from geographically isolated habitats, such as islands. One such barrier in Brazil is the highly isolated combined oceanic reefs of Atol das Rocas (AR) and Fernando de Noronha (FDN; Floeter et al. 2008; Rocha 2003; Rocha, Robertson, Roman, & Bowen, 2005). Although many populations are able to sustain connectivity across the 260 km separating AR from the Brazilian coast, 5% of fishes are endemic, which suggests that some taxa are unable to consistently exchange genes with populations on the coast (Floeter et al. 2008). As a recent example, a new species of goby (*Bathygobius brasiliensis*) was described that is restricted to AR and FDN (Rodríguez-Rey, Filho, Araújo, & Solé-cava, 2017).

Considering how oceanographic features affect evolutionary history and population connectivity, low dispersal organisms are ideal to evaluate the impact of permeable barriers in the western Atlantic. The bridled goby (*Coryphopterus glaucofraenum*) and sand-canyon goby (*C. venezuelae*) are small (<55 mm), sedentary, benthic fishes that territorially defend nests located on sandy patches near coral reefs (Forrester, Harmon, Helyer, Holden, & Karis, 2010). Early studies of *C. glaucofraenum* described individuals as having morphological variation but not enough to be considered multiple species (Böhlke & Robins, 1960). Subsequently, *C. venezuelae* was elevated to full species designation based on the number of fin elements, slight pigmentation patterns and genetic differentiation using cytochrome oxidase I (COI; Baldwin, Weigt, Smith, & Mounts, 2009). Both species, occur throughout the entire Caribbean while *C. glaucofraenum* extends to southern Brazil

(Robins and Ray 1986). However, because most gobies are demersal spawners and often demonstrate significant genetic structure across a wide range (Milá, Van Tassell, Calderón, Rüber, & Zardoya, 2017), it is likely that *C. glaucofraenum* and *C. venezuelae* are impacted by dispersal barriers throughout the western Atlantic.

Due to sparse sampling from across their range, low dispersal potential, and the presence of several barriers, I used sampling across a broad geographic scale to test if species- or population-level differences were present within *C. glaucofraenum* and *C. venezuelae*. First, I hypothesized that two major barriers, the Amazon River outflow and the isolated Brazilian island, would harbor unique genetic clades indicative of species-level genetic divergence. This would result in four monophyletic clades including *C. venezuelae* and three *C. glaucofraenum*: a Caribbean, Brazilian and a Brazilian island lineage. Second, I hypothesized that minor barriers in each province would promote population structure within these clades. In accordance with one of the most prominent Caribbean trends, populations within the eastern or western Caribbean should be more genetically similar to each other than populations compared across the Caribbean, while the Bahamas often show distinctness from either area (DeBiasse et al., 2016; Foster et al., 2012; Jackson et al., 2014; Taylor & Hellberg, 2003, 2006). Similarly, coastal Brazilian populations should be separated into three genetic clusters based on the SEC and Cabo Frio barriers resulting in northern, central and southern clusters (Boschi, 2000; Cunha et al., 2014; Fernandes et al., 2012; Machado et al., 2017; Maggioni et al., 2003; Picciani et al., 2016; Santos et al., 2006). Additionally, I sought to infer demographic events, such as migration and bottlenecks, that often result in reduced genetic diversity (Cornuet & Luikart, 1996). Based on the fact that

most of the species in the genus *Coryphopterus* occur in the Caribbean (Baldwin et al., 2009), I hypothesized that migration occurred in a southward direction across the Amazon barrier and continued south once established in Brazil. Therefore, I assessed genetic diversity among clades and among populations within clades to check for evidence of bottlenecks. Lastly, these results are discussed as they relate to broader phylogeographic trends of marine taxa in the western Atlantic.

CHAPTER TWO: METHODS

COI Sequence Variation and Phylogenetics

I analyzed 112 individuals of *C. glaucofraenum* collected throughout the Brazilian coast and supplemented these with 94 individuals of *C. glaucofraenum* and *C. venezuelae* from the Caribbean through GenBank (Figure 1, Table 1). Tissue and fin clips from field capture were immediately placed in 95% ethanol and subsequently frozen for long-term storage. I extracted genomic DNA using a Serapure bead protocol (Rohland & Reich, 2012) and amplified the COI gene with FishF1 and FishR1 primers (Ward, Zemlak, Innes, Last, & Hebert, 2005). PCR was held in 20 µl reactions using 1-10 ng genomic DNA, 2 µl 10x buffer, 0.8mM DNTPs, 1.63mM $MgCl_2$, 0.5µM forward and reverse primers and 0.2 µl Taq DNA polymerase. Thermal cycling parameters consisted of a 4 minute denaturation at 94 °C followed by 35 cycles of 30 second denaturation at 94 °C, 35 seconds of annealing at 53 °C, 45 seconds of extension at 72 °C and a final extension of 7 minutes at 72 °C. PCR products were sent to Eurofins Genomics for sequencing. Following sequencing, I verified chromatographs by eye using Sequencher 5.1 (Gene Codes, Ann Arbor, MI, USA). Sequences were then trimmed and aligned with GenBank samples using MEGA7 (Kumar, Stecher, & Tamura, 2016) followed by file formatting for each analysis using PGDSpider (Lischer & Excoffier, 2012).

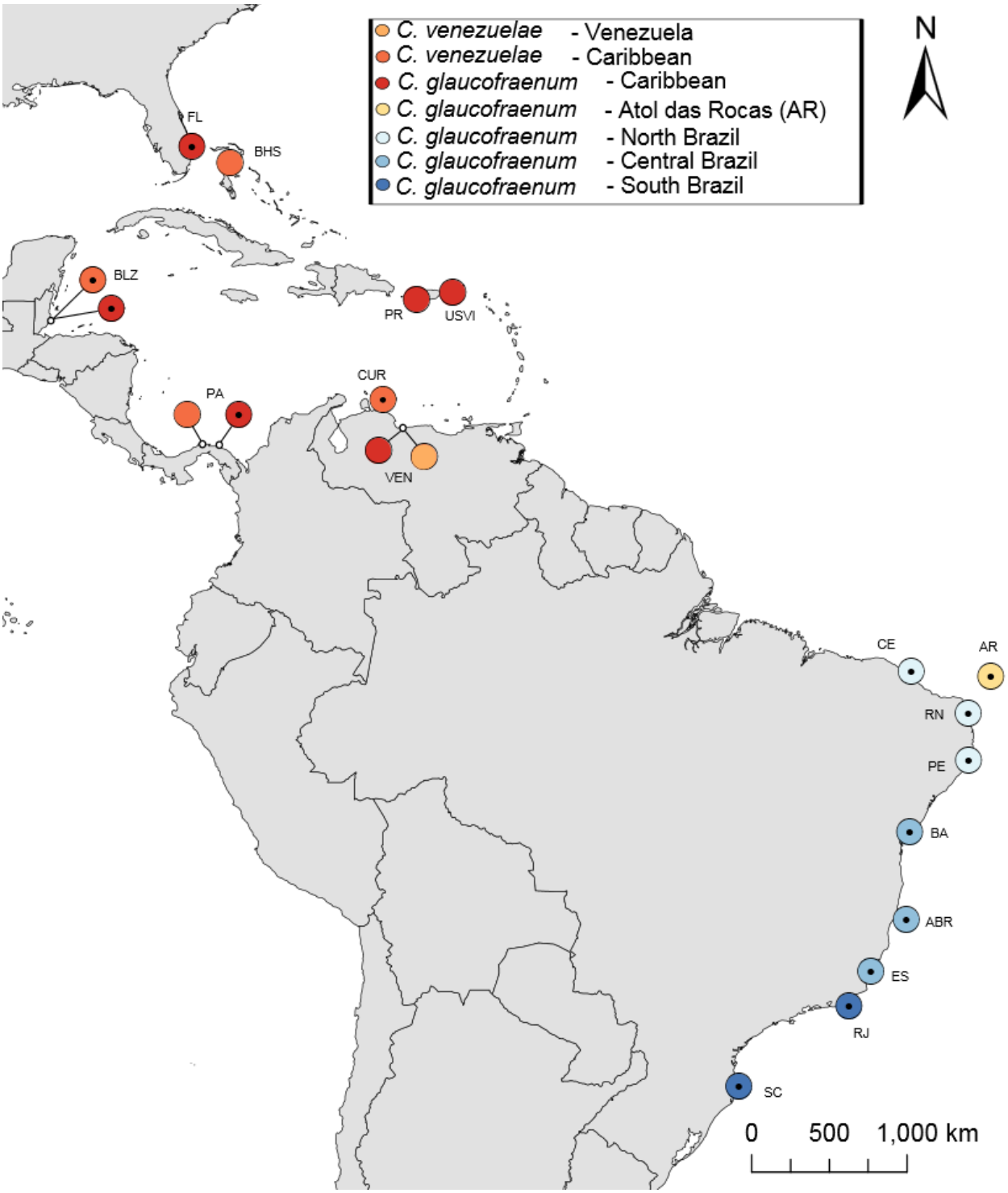

Figure 1. Map of study area in the western Atlantic. All sampling points include COI data while circles with dots indicate that SNPs were also used.

Table 1. Collection location and genetic diversity estimates for COI and SNP datasets with standard deviation in parentheses. Populations with < 5 samples were not included in estimates of genetic diversity. Distinct clades are in bold while population clusters are italicized. Number of samples (N); number of haplotypes (N_h); haplotype diversity (h); nucleotide diversity (π); number of effective alleles (N_a); observed heterozygosity (H_O); unbiased expected heterozygosity (uH_e).

	Cytochrome oxidase I				SNPs			
Location	N	N_h	h	π	N	N_a	H_O	uH_E
Caribbean (*C. venezuelae*)	39	19	0.897 (0.032)	5.74×10^{-3} (0.42×10^{-3})	11	1.150	0.077	0.096
Bahamas (CVEN BHS)	3	3	1.0 (0.272)	6.01×10^{-3} (1.7×10^{-3})	-	-	-	-
Belize (CVEN BLZ)	8	6	0.929 (0.084)	4.25×10^{-3} (0.73×10^{-3})	5	1.131	0.078	0.090
Panama (CVEN PA)	5	4	0.900 (0.161)	5.05×10^{-3} (1.38×10^{-3})	-	-	-	-
Venezuela (CVEN VEN)	13	4	0.423 (0.164)	0.83×10^{-3} (0.36×10^{-3})	-	-	-	-
Curacao (CVEN CUR)	10	7	0.911 (0.077)	3.0×10^{-3} (0.62×10^{-3})	6	1.136	0.075	0.091
Caribbean (*C. glaucofraenum*)	55	19	0.660 (0.074)	2.50×10^{-3} (0.48×10^{-3})	20	1.206	0.098	0.128
Florida (FL)	6	3	0.600 (0.215)	1.65×10^{-3} (0.74×10^{-3})	5	1.183	0.109	0.124
US Virgin Islands (USVI)	5	3	0.700 (0.218)	4.28×10^{-3} (1.25×10^{-3})	-	-	-	-
Puerto Rico (PR)	2	2	1.000 (0.500)	13.18×10^{-3} (6.59×10^{-3})	-	-	-	-
Belize (BLZ)	17	7	0.596 (0.139)	1.74×10^{-3} (0.60×10^{-3})	7	1.188	0.098	0.121
Panama (PA)	16	3	0.425 (0.133)	0.95×10^{-3} (0.38×10^{-3})	8	1.196	0.092	0.126
Venezuela (VEN)	9	5	0.806 (0.120)	2.75×10^{-3} (0.88×10^{-3})	-	-	-	-
Atol das Rocas (AR)	9	5	0.861 (0.087)	3.37×10^{-3} (1.0×10^{-3})	2	-	-	-
Brazil	103	31	0.698 (0.0003)	2.34×10^{-3} (0.27×10^{-3})	55	1.162	0.081	0.102
North Brazil					18	1.156	0.085	0.100
Ceará (CE)	4	2	0.500 (0.265)	0.89×10^{-3} (0.47×10^{-3})	1	-	-	-
Rio Grande de Norte (RN)	9	6	0.833 (0.127)	2.77×10^{-3} (0.75×10^{-3})	9	1.157	0.096	0.102
Pernambuco (PE)	15	7	0.819 (0.082)	2.44×10^{-3} (0.48×10^{-3})	8	1.150	0.075	0.098
Central Brazil					22	1.158	0.081	0.100
Bahia (BA)	15	6	0.705 (0.114)	2.65×10^{-3} (0.53×10^{-3})	6	1.141	0.059	0.094
Abrolhos (ABR)	12	9	0.909 (0.079)	2.92×10^{-3} (0.66×10^{-3})	9	1.157	0.089	0.102
Espirito Santo (ES)	18	10	0.869 (0.059)	3.53×10^{-3} (0.40×10^{-3})	7	1.151	0.089	0.099
South Brazil					15	1.140	0.080	0.096
Rio de Janeiro (RJ)	14	2	0.143 (0.119)	0.51×10^{-3} (0.20×10^{-3})	8	1.148	0.077	0.096
Santa Catarina (SC)	16	3	0.342 (0.140)	0.64×10^{-3} (0.28×10^{-3})	7	1.149	0.075	0.097

In order to determine the evolutionary relationships among lineages, I performed a Bayesian phylogenetic analysis using Beast2 (Bouckaert et al., 2014) with the HKY+G model of evolution as determined in PartitionFinder (Lanfear, Calcott, Ho, & Guindon, 2012). I used a related species, *C. tortugae*, as an outgroup and performed four independent runs of 100 million generations each with samples being taken every 10,000 generations. Each run was checked in Tracer v 1.6 (Bouckaert et al., 2014) to ensure effective sample sizes (ESS) were ≥200 for each parameter. LogCombiner v 2.4.7 (Bouckaert et al. 2014) was used to discard 10% burnin for each run and combine a subset of trees from each run for a total of 9000 tree states. Using this combined file, I used TreeAnnotator to create a 50% majority-rule consensus tree which was viewed in FigTree v 1.4.2 (Rambaut, 2016).

I verified clades found from phylogenetic analysis using three approaches. First, I followed Baldwin et al. (2009) by evaluating genetic distance between clades to see if clade divergence indicates species-level differences. Pairwise distances between clades were calculated in MEGA7 using the Kimura 2-parameter model. Second, I created a TCS (Clement, Posada, & Crandall, 2000) haplotype network in PopART (Leigh & Bryant, 2015) to visualize the distribution of haplotypes among clades and populations. Lastly, to verify genetic partitioning among clades found in the phylogenetic analysis, I performed an analysis of molecular variance (AMOVA) in Arlequin 3.5.2 (Excoffier & Lischer, 2010) using the Tamura and Nei (1993) distance method and 20,000 permutations.

COI Variation within Clades

I estimated pairwise ϕ_{ST}, an analog of F_{ST}, among all populations using 20,000 permutations with the Tamura and Nei (1993) substitution model to determine population differentiation. Within each clade, I expected populations within an area to be more similar to each other than populations across a barrier. As a result, I compared pairwise ϕ_{ST} in *C. glaucofraenum* between populations within the east (USVI, PR, VEN, CUR) and west (FL, BLZ, PA) Caribbean to pairwise ϕ_{ST} between east-west population pairs within each clade using a student's t-test in R studio (R Core Team, 2013), but sparse sampling prohibited a similar analysis for *C. venezuelae*. To evaluate whether barriers impact regional connectivity in Brazil, I also tested whether populations within northern, central, or southern Brazil were more closely related to each other than population pairs across these regions using a Wilcoxon rank-sum test.

Tajima's *D* was calculated in DNAsp v5 (Librado & Rozas, 2009) to test if clades identified in the phylogenetic analysis show evidence of demographic expansion or contraction. Here, negative values of Tajima's D indicate population expansion and positive values suggest populations have recently contracted. I also tested to see if genetic diversity differed among clades due to bottleneck events, so I measured haplotype and nucleotide diversities for each population. Levels of diversity were then compared among clades with a non-parametric Kruskal-Wallis test. Similarly, to see if populations in southern Brazil demonstrated lower diversity due to a recent founder event, diversity among the north, central and southern populations in Brazil were compared using a Kruskal- Wallis test.

SNP Generation and Filtering

A reduced-sample SNP dataset was generated using 103 individuals from 16 populations across the range of both *C. glaucofraenum* and *C. venezuelae* including three individuals of *C. tortugae* as an outgroup (Figure 1, Table 1). Genomic DNA was converted into nextRAD genotyping-by-sequencing libraries (SNPsaurus, LLC) as in Russello, Waterhouse, Etter, and Johnson (2015). Briefly, genomic DNA was first fragmented with Nextera reagent (Illumina, Inc), which also ligates short adapter sequences to the ends of fragments. The Nextera reaction was scaled for fragmenting 7 ng of genomic DNA, although 17.5 ng of genomic DNA was used for input to compensate for the amount of degraded DNA in the samples and to increase fragment sizes. Fragmented DNA was then amplified for 26 cycles at 73 °C, with one of the primers matching the adapter and extending 9 nucleotides into the genomic DNA with the selective sequence GTGTAGAGG. Thus, only fragments starting with a sequence that can be hybridized by the selective sequence of the primer will be efficiently amplified. The nextRAD libraries were sequenced on an Illumina HiSeq 4000 with one lane of 150 bp reads (University of Oregon).

Genotyping analysis used custom scripts (SNPsaurus, LLC) that trimmed the reads using bbduk (BBMap tools). Next, a *de novo* reference was created by collecting 10 million reads in total, evenly from the samples. To account for potential paralogs, *de novo* reference excluded reads that had counts fewer than 10 or more than 1,000. The remaining loci were then aligned to each other to identify allelic loci and collapse allelic haplotypes to a single representative. All reads were mapped to the reference with an alignment identity

threshold of 95% using bbmap (BBMap tools). Genotype calling was completed using SAMtools and BCFtools (Li, 2011; Li et al., 2009), followed by filtering to remove alleles with a population frequency of less than 3%. Loci were removed that were heterozygous in all samples or had more than 2 alleles in a sample suggesting collapsed paralogs. The absence of artifacts was checked by counting SNPs at each read nucleotide position and verifying that SNP number did not increase with reduced base quality at the end of the read. Additional filtering using VCFtools (Danecek et al., 2011) removed loci that had less than 10x coverage, minor alleles with a frequency of less than 0.05 and any sites with >20% missing data. After previous filtering was completed, the dataset was thinned to keep only one SNP per fragment to reduce linkage between loci. All remaining loci were evaluated for Hardy-Weinberg equilibrium (HWE) so that loci were removed if more than seven populations were out of HWE at $p = 0.01$. Any individual with >20% missing data was excluded from analyses.

SNP Phylogenetics

To estimate evolutionary relationships among species, I used three phylogenetic methods. First, I utilized a Bayesian approach in MrBayes v3.2.6 (Ronquist & Huelsenbeck, 2003) through the Cipres Science Gateway (Miller et al. 2010). Two independent runs were performed with four chains for a total of 30 million generations with sampling taken every 10,000 generations and a 25% burnin. Using jModelTest2 v2.1.10 (Darriba, Taboada, Doallo, & Posada, 2012), the GTR+G model of evolution was used based on the corrected Akaike's Information Criterion (AICc). Due to the large amount of missing data in some

samples (see Results), each clade was constrained to monophyly in order to accurately assess the relationships among species. Constraining these taxa is justified based on the strong support of the COI dataset (see Results). By constraining several taxa, I am still able to infer relationships among and within clades. Second, I performed a maximum likelihood analysis in RAxML (Stamatakis, 2014) through the Cipres Science Gateway. All loci were concatenated and a correction bias (Lewis, 2001) was implemented due to using all variable sites. A GTR + G nucleotide substitution model was implemented followed by 1,000 bootstraps for likelihood estimation. Similar to the Bayesian approach, I constrained the two samples from AR to monophyly due to the large amount of missing data and allowed all other taxa to remain unconstrained. Lastly, I utilized a fast, coalescent-based approach with SVDquartets (Chifman & Kubatko, 2014) implemented in Paup (Swofford 2002) which first estimates gene trees, then infers a species tree. All trees were visualized and modified in FigTree.

SNP Population Genetics

To see if population structure existed within clades identified in the phylogenetic analysis, a Bayesian clustering analysis was performed within each clade and without population location priors in STRUCTURE (Pritchard, Stephens, & Donnelly, 2000). Here, I performed ten runs for each population (K) up to the maximum number of populations within each clade using a 50,000-replicate burnin and 500,000 replicates for each run. The Evanno ΔK method (Evanno, Regnaut, & Goudet, 2005) was used in StructureHarvester (Earl & VonHoldt, 2012) to determine the most likely value for K. After initial runs were

complete, I checked for substructure by rerunning STRUCTURE within genetic clusters using the same parameters. Because STRUCTURE analyses did not always provide clear patterns, I also tested for genetic partitioning within clades using AMOVA. First, I evaluated putative barriers within Brazil by partitioning populations into north, central or south Brazil. As a control, I incorrectly grouped populations during a single AMOVA to show that no variation was explained among groups. Therefore, a correct clustering of populations should increase the amount of variation explained relative to the control. To see if minor barriers in Brazil explain genetic clustering, I then compared two alternative AMOVAs that showed either two (north-central and south) or three groups (north, central and south). The AMOVA with the most variation explained was considered to be the more likely clustering of populations. All AMOVAs were implemented in Arlequin with 20,000 permutations.

To determine population differentiation, I estimated pairwise F_{ST} among populations using the pairwise distance approach in Arlequin with 20,000 permutations. Following the same approach as with COI data, I then compared levels of F_{ST} between population pairs from the same area to levels of F_{ST} between population pairs from different areas using a student's t-test in R studio (R Core Team, 2013). However, there were too few population pairs within the east and west Caribbean for *C. glaucofraenum* and *C. venezuelae*, so I exclusively analyzed Brazilian populations using this approach. Specifically, I tested to see if populations within north, central or southern Brazil were more similar to each other than population pairs from different areas of Brazil. Additionally, isolation-by-distance (IBD) was tested in GenePop (Rousset, 2008) to see if populations were dispersal limited within Brazil.

Lastly, to see if genetic diversity varied among clades due to a potential bottleneck or founder event, I compared levels of genetic diversity among clades and among populations within clades. I estimated the number of effective alleles, observed heterozygosity, and expected heterozygosity to approximate overall genetic diversity. Then, expected heterozygosity was compared among populations using a Kruskal-Wallace test to see if populations exhibit signs of bottleneck in the form of low genetic diversity. Genetic diversity was not estimated at locations with < 5 individuals due to inaccuracy with small population sizes (Nazareno, Bemmels, Dick, & Lohmann, 2017).

CHAPTER THREE: RESULTS

COI Phylogenetics

Using a 690 bp alignment of COI, I found four highly supported monophyletic clades, all four are highly divergent from one another (Figure 2, Table 2). While the most basal node showed poor support, most other nodes exhibited high support (> 0.95 posterior support). Even though Brazil and AR are close in proximity, AR appeared to be more closely related to *C. venezuelae* than to Brazil. AR and Brazil were previously described as *C. glaucofraenum*, although there was strong posterior support to suggest that Brazil and AR are more closely related to *C. venezuelae* than either clade is to *C. glaucofraenum* (Figure 2). Because the current taxonomy is paraphyletic, samples from the Caribbean *C. glaucofraenum* clade will be referred to as *C. glaucofraenum*, while samples from the Brazilian coast and offshore island (i.e. Atol das Rocas) will be referenced as Brazil and AR. The AMOVA performed on all four clades corroborates the distinction of each taxon with 96% of variation in the data explained among taxa ($p < 0.001$; Table 3). Moreover, percent sequence divergence among the four primary clades ranged from 6.58% (between *C. venezuelae* and AR) to 13.29% (between *C. glaucofraenum* and Brazil; Table 2). Within *C. venezuelae*, individuals collected from Venezuela showed strong support for monophyly despite having diverged <1% from the rest of the Caribbean *C. venezuelae* samples (Table 2). The haplotype network revealed that none of the 74 haplotypes were shared among clades and a minimum of 27 (AR-*C. venezuelae*) and a maximum of 45 mutations (*C. venezuelae*-*C. glaucofraenum* and AR-*C. glaucofraenum*) connected haplotypes between

clades (Figure 3). The overall star-shape configuration of the haplotype network suggests Brazil and *C. glaucofraenum* have undergone a recent expansion. As with the phylogeny above, the haplotype network showed that within *C. venezuelae*, the Venezuela population was at least two mutations away from any other Caribbean population. In contrast, Brazilian haplotypes were evenly distributed among areas with no significant structure detected across barriers.

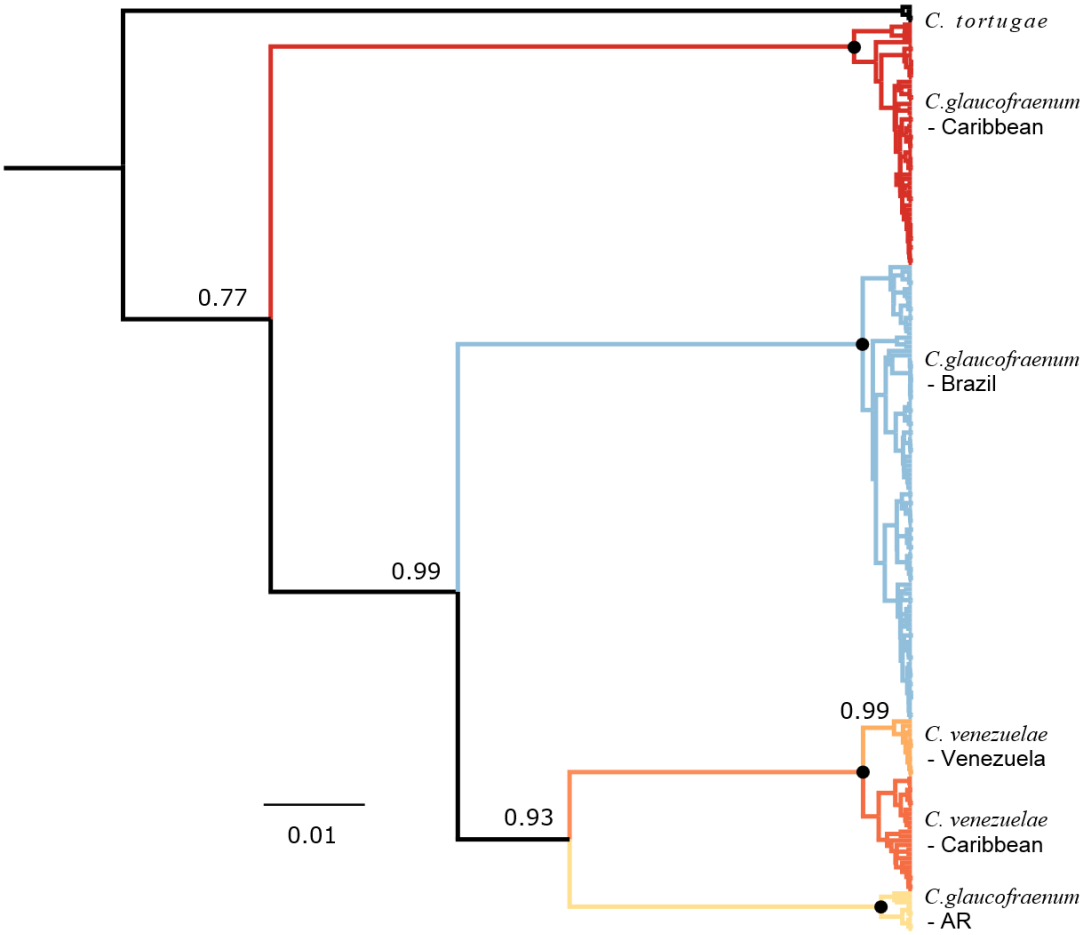

Figure 2. Bayesian phylogeny of COI. Clades are colored as in collection sites and posterior support values are shown at nodes with black circles representing a posterior value of one.

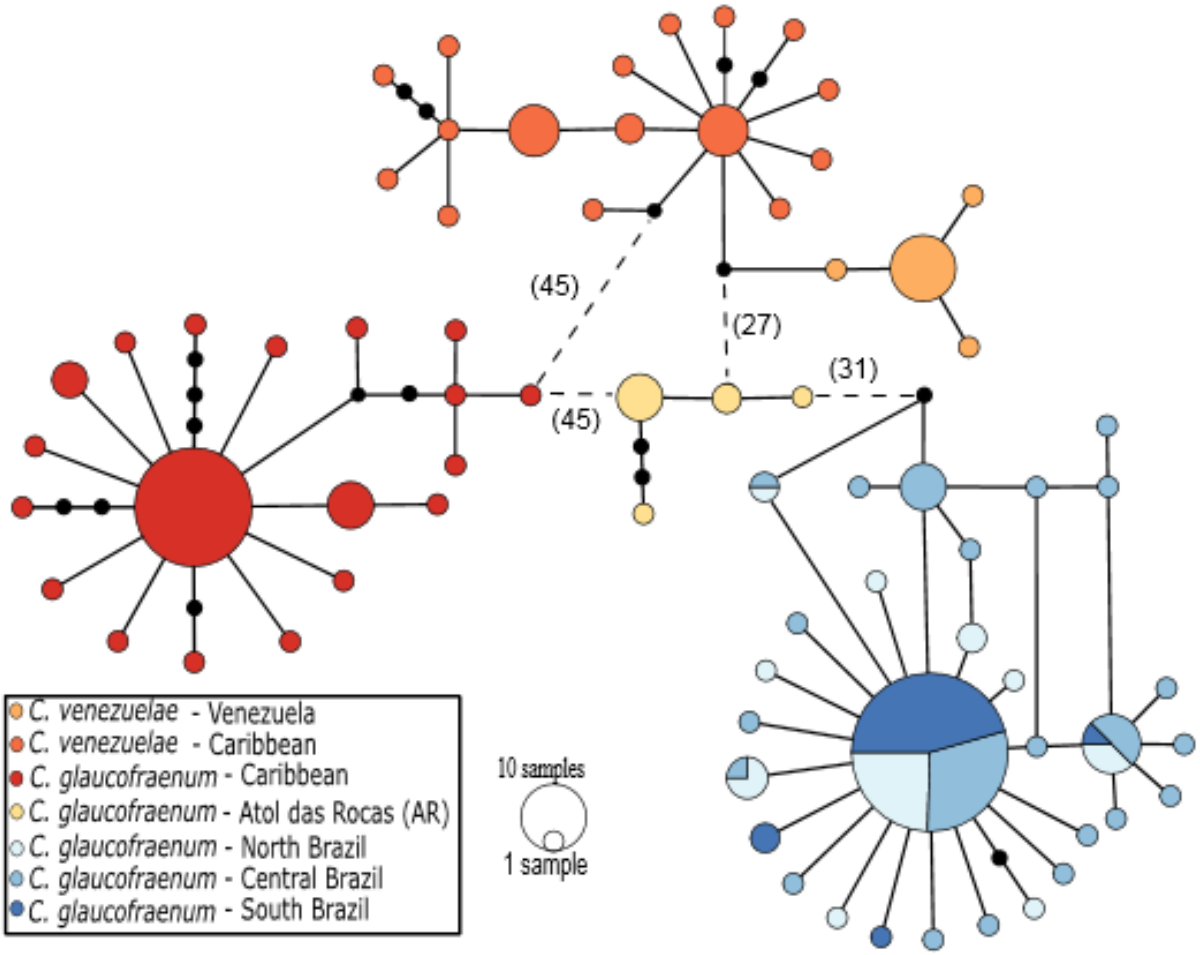

Figure 3. Haplotype network of COI data. Colors correspond to lineages while shading represents population differentiation. Dashed lines with numbers indicate the number of inferred mutations between lineages.

COI Population Genetics

Most ϕ_{ST} estimates among populations in different clades were high (>0.91) and significantly different from zero, which indicates a lack of gene flow among clades (Table 4). Conversely, populations within clades shared more gene flow as indicated by their smaller ϕ_{ST} estimates (-0.11 – 0.78; Table 4).

Table 2. Percent sequence divergence between and within taxa (bold) using the pairwise Kimura 2-parameter model.

	C. dicrus	*C. glaucofraenum*	*C. venezuelae*	VEN *C. venezuelae*	AR	Brazil
C. dicrus	-					
C. glaucofraenum	24.44%	**0.28%**				
C. venezuelae	21.57%	11.42%	**0.43%**			
VEN *C. venezuelae*	21.12%	11.98%	0.88%	**0.09%**		
AR	22.91%	11.50%	6.58%	6.51%	**0.36%**	
Brazil	24.53%	13.29%	7.82%	7.30%	7.62%	**0.24%**

Table 3. a) AMOVA results for COI data that partitioned four and five groups. b) AMOVA results for SNP data that partitioned areas of Brazil.

	Sum of squares	Variance Components	Variation	p-value
a) COI				
Four groups: *C. glaucofraenum, C. venezuelae*, AR, Brazil				
Among Groups	3361.86	25.04	96.41	<0.001
Among Populations within Groups	55.47	0.28	1.07	<0.001
Within Populations	123.17	0.66	2.52	<0.001
Five groups: *C. glaucofraenum, C. venezuelae*, Venezuela, AR, Brazil				
Among Groups	3394.51	24.66	97.01	<0.001
Among Populations within Groups	25.82	0.11	0.42	<0.001
Within Populations	123.17	0.66	2.58	<0.001
b) SNPs				
Control- Brazil: North, Central/South				
Among Groups	241.19	1.79	1.65	0.054
Among Populations within Groups	874.68	3.25	3.01	<0.001
Within Populations	10519.45	103.13	95.34	<0.001
Two groups- Brazil: North/Central, South				
Among Groups	321.31	4.23	3.86	0.036
Among Populations within Groups	794.56	2.22	2.02	<0.001
Within Populations	10519.45	103.13	94.12	<0.001
Three groups- Brazil: North, Central, South				
Among Groups	524.14	3.88	3.58	0.004
Among Populations within Groups	591.72	1.20	1.1	0.416
Within Populations	10519.45	103.13	95.31	<0.001

Populations of *C. glaucofraenum* were more similar if they were in the same area (i.e. within east or within west Caribbean) as opposed to populations from different areas (i.e. east vs. west comparisons; $t = -2.44$, $df = 13$, $p = 0.01$).

Table 4. Pairwise ϕ_{ST} (lower) and F_{ST} (upper) estimates among all locations. Values in bold are significant at p <0.05 for ϕ_{ST} and p<0.01 for F_{ST} after 20,000 permutations.

	FL	USVI	PR	BLZ	PA	VEN	CVEN BHS	CVEN BLZ	CVEN PA	CVEN VEN	CVEN CUR	AR	CE	RN	PE	BA	ABR	ES	RJ	SC
FL		-	-	**0.04**	**0.04**	-	-	**0.74**	-	-	**0.74**	0.50	0.78	**0.78**	**0.78**	**0.78**	**0.78**	**0.78**	**0.78**	**0.78**
USVI	0.16		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PR	**0.38**	0.02		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
BLZ	0.00	**0.21**	**0.51**		**0.04**	-	-	**0.74**	-	-	**0.74**	0.48	0.78	**0.78**	**0.79**	**0.78**	**0.78**	**0.78**	**0.79**	**0.79**
PA	0.09	**0.35**	**0.66**	0.03		-	-	**0.74**	-	-	**0.75**	0.49	0.78	**0.78**	**0.78**	**0.78**	**0.78**	**0.78**	**0.78**	**0.78**
VEN	**0.19**	**0.20**	**0.40**	**0.21**	**0.31**		-	-	-	-	-	-	-	-	-	-	-	-	-	-
CVEN BHS	**0.97**	**0.95**	0.91	**0.98**	**0.99**	**0.97**		-	-	-	-	-	-	-	-	-	-	-	-	-
CVEN BLZ	**0.96**	**0.95**	**0.94**	**0.97**	**0.98**	**0.96**	0.05		-	-	**0.09**	0.43	0.78	**0.74**	**0.75**	**0.76**	**0.75**	**0.74**	**0.75**	**0.75**
CVEN PA	**0.97**	**0.95**	**0.93**	**0.97**	**0.98**	**0.96**	-0.11	-0.07		-	-	-	-	-	-	-	-	-	-	-
CVEN VEN	**0.99**	**0.98**	**0.98**	**0.99**	**0.99**	**0.98**	**0.79**	**0.73**	**0.78**		-	-	-	-	-	-	-	-	-	-
CVEN CUR	**0.97**	**0.96**	**0.96**	**0.98**	**0.98**	**0.97**	0.01	**0.20**	0.11	**0.75**		0.40	0.77	**0.74**	**0.75**	**0.75**	**0.74**	**0.74**	**0.75**	**0.75**
AR	**0.98**	**0.97**	**0.96**	**0.98**	**0.99**	**0.97**	**0.94**	**0.93**	**0.94**	**0.97**	**0.95**		0.66	0.28	0.29	0.35	0.32	0.29	0.29	0.31
CE	**0.99**	**0.97**	0.96	**0.98**	**0.99**	**0.98**	**0.96**	**0.94**	**0.95**	**0.99**	**0.96**	**0.97**		0.03	0.02	0.11	0.09	0.04	0.08	0.10
RN	**0.98**	**0.97**	**0.96**	**0.98**	**0.99**	**0.98**	**0.95**	**0.94**	**0.95**	**0.98**	**0.96**	**0.96**	-0.07		0.00	**0.03**	**0.02**	**0.02**	**0.05**	**0.06**
PE	**0.98**	**0.98**	**0.97**	**0.98**	**0.99**	**0.98**	**0.96**	**0.95**	**0.96**	**0.98**	**0.96**	**0.97**	-0.02	-0.01		**0.04**	**0.03**	**0.03**	**0.06**	**0.07**
BA	**0.98**	**0.97**	**0.97**	**0.98**	**0.98**	**0.98**	**0.96**	**0.95**	**0.95**	**0.97**	**0.96**	**0.96**	0.07	0.03	0.04		0.03	0.01	**0.06**	**0.07**
ABR	**0.98**	**0.97**	**0.97**	**0.98**	**0.98**	**0.97**	**0.95**	**0.94**	**0.95**	**0.97**	**0.96**	**0.96**	-0.07	-0.03	-0.03	0.03		0.00	**0.05**	**0.05**
ES	**0.97**	**0.97**	**0.96**	**0.98**	**0.98**	**0.97**	**0.95**	**0.94**	**0.95**	**0.96**	**0.95**	**0.95**	0.13	**0.13**	**0.16**	**0.13**	**0.13**		**0.05**	**0.06**
RJ	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	**0.98**	**0.97**	**0.98**	**0.99**	**0.98**	**0.98**	0.06	-0.01	0.01	0.09	-0.02	**0.21**		0.02
SC	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	**0.98**	**0.97**	**0.98**	**0.99**	**0.98**	**0.98**	0.06	0.05	**0.07**	**0.18**	**0.04**	**0.26**	0.02	

Considering specific pairwise ϕ_{ST} comparisons between *C. glaucofraenum* populations, Venezuela was highly isolated from all Caribbean populations (ϕ_{ST} = 0.19 – 0.40), and with only two unique haplotypes, Puerto Rico was significantly isolated from all populations (ϕ_{ST} = 0.38 – 0.66) except the nearby US Virgin Islands (ϕ_{ST} = 0.02). Similar isolation of Venezuela was found in *C. venezuelae* including strong isolation from Curaçao (ϕ_{ST} = 0.75), located only 231 km away. In contrast, most other Caribbean populations were genetically similar despite much longer distances between sites (ϕ_{ST} = -0.11 - 0.20; Table 4, Figure 1). In Brazil, the only populations that demonstrated significant levels of differentiation were Santa Catarina (ϕ_{ST} = 0.02 - 0.18) and Espírito Santo (ϕ_{ST} = 0.13 - 0.26). In addition, when testing to see if *a priori* Brazilian populations in the same area (north, central or south) had lower ϕ_{ST} estimates than pairwise populations across a barrier, there were no differences detected (Wilcoxon rank-sum test; W = 50.5, p-value = 0.11).

Haplotype diversity ranged widely from 1.0 in Puerto Rico and the Bahamas to 0.143 in Rio de Janeiro while nucleotide diversity ranged from 13.18 x 10^{-3} to 0.51 x 10^{-3} in the same populations (Table 1). Comparing diversity among the four clades, haplotype (Kruskal-Wallis χ^2 = 3.51, df = 3, p =0.32) and nucleotide diversities (Kruskal-Wallis χ^2 = 3.65, df = 3, p = 0.30) were not significantly different. Furthermore, comparing among north, central and southern Brazil, haplotype (Kruskal-Wallis χ^2 = 4.70, df = 2, p-value = 0.10) and nucleotide (Kruskal-Wallis χ^2 = 5.36, df = 2, p-value = 0.07) diversities were not significantly different among areas. Tajima's D was significantly negative for both Brazil (*D* = -2.22, p < 0.01) and *C. glaucofraenum* (*D* = -2.39, p < 0.01) indicating that each clade had

undergone a recent population expansion, as was suggested above by the shape of the haplotype network. However, Tajima's D in *C. venezuelae* (D = -1.36, $p > 0.10$) and AR (D = -1.04, $p > 0.10$) were not significantly different from zero.

SNP Filtering

Although 103 individuals were sent out for SNP genotyping, the final data set included 91 samples because 12 were removed due to poor quality sequencing. After *de novo* assembly and initial filtering, there was a total of 9,003 SNPs. Following additional locus filtration for 10x coverage, 20% missing data and HWE, samples were thinned to include only one SNP per fragment resulting in a final dataset of 2,401 SNPs. Despite falling below the *a priori* threshold for missing data within an individual, the two samples from AR were maintained in the dataset due to their importance for phylogenetic analyses.

SNP Phylogenetics

As with the COI tree above, the Bayesian tree showed strong support for four monophyletic clades (Figure 4). There was high support for the overall clade consisting of *C. venezuelae*, AR and Brazil with strong support for AR and Brazil being sister taxa. Similarly, *C. venezuelae* nodes were strongly supported, particularly for Belizean individuals, which formed a monophyletic clade (Figure 4). In contrast, samples within the Brazilian clade largely consisted of a polytomy. The maximum likelihood (Figure 4) and coalescent analyses (Figure 5) were topologically identical to the Bayesian tree. The same relationships among major clades were recovered and each clade exhibited high bootstrap support for monophyly. Similar to the Bayesian tree, maximum likelihood analysis found all

C. venezuelae individuals from Belize to be monophyletic and showed moderate bootstrap support.

SNP Population Genetics

Using the Δ*K* approach, Bayesian clustering analyses in STRUCTURE indicated *K* = 2in *C. glaucofraenum*, though the split does not conform to any particular location (Figure 6). Similarly, *C. venezuelae* individuals clustered into *K* = 2 which also did not appear to match any known barriers. Despite this, two individuals from Curaçao (*C. venezuelae*) were strongly differentiated from the remainder of the Caribbean individuals (Figure 6).

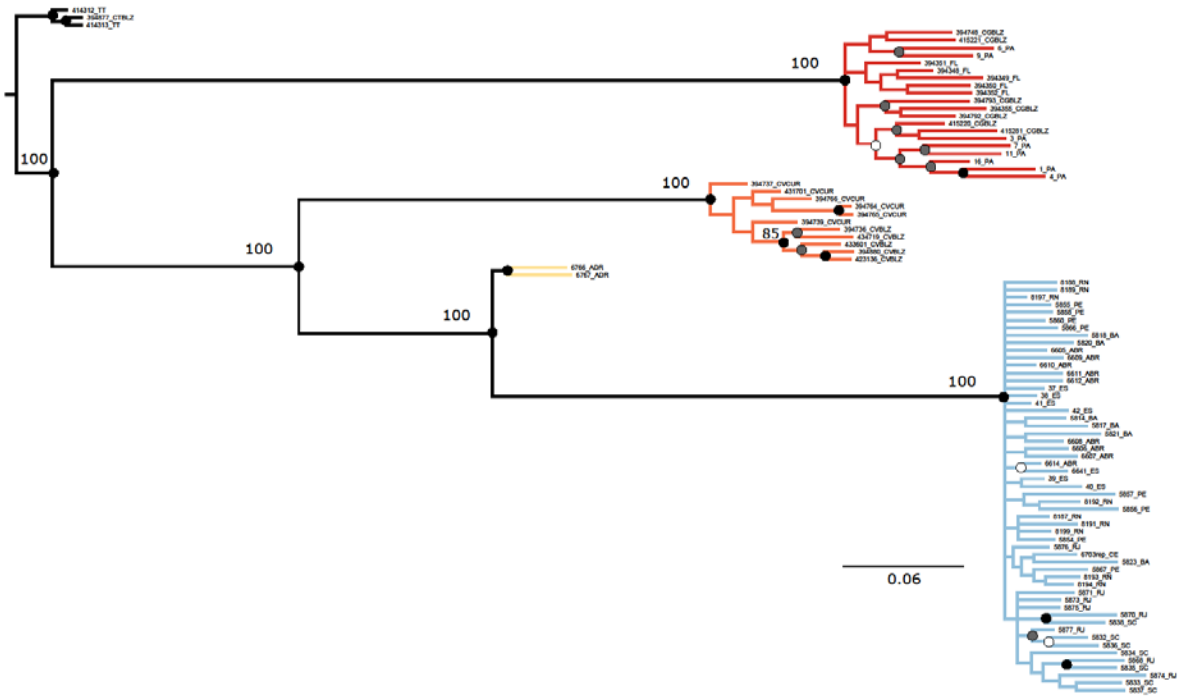

Figure 4. Bayesian phylogeny of SNP data. Black, grey and white nodes represent posterior probabilities of one, ≥0.95 and ≥0.90 respectively. Values above nodes represent bootstraps from clades found in RAxML analysis. Each clade was constrained to monohphyly for the MrBayes tree while only the two samples from AR were constrained in the RAxML tree. This approach resulted in an identical topology between the two approaches.

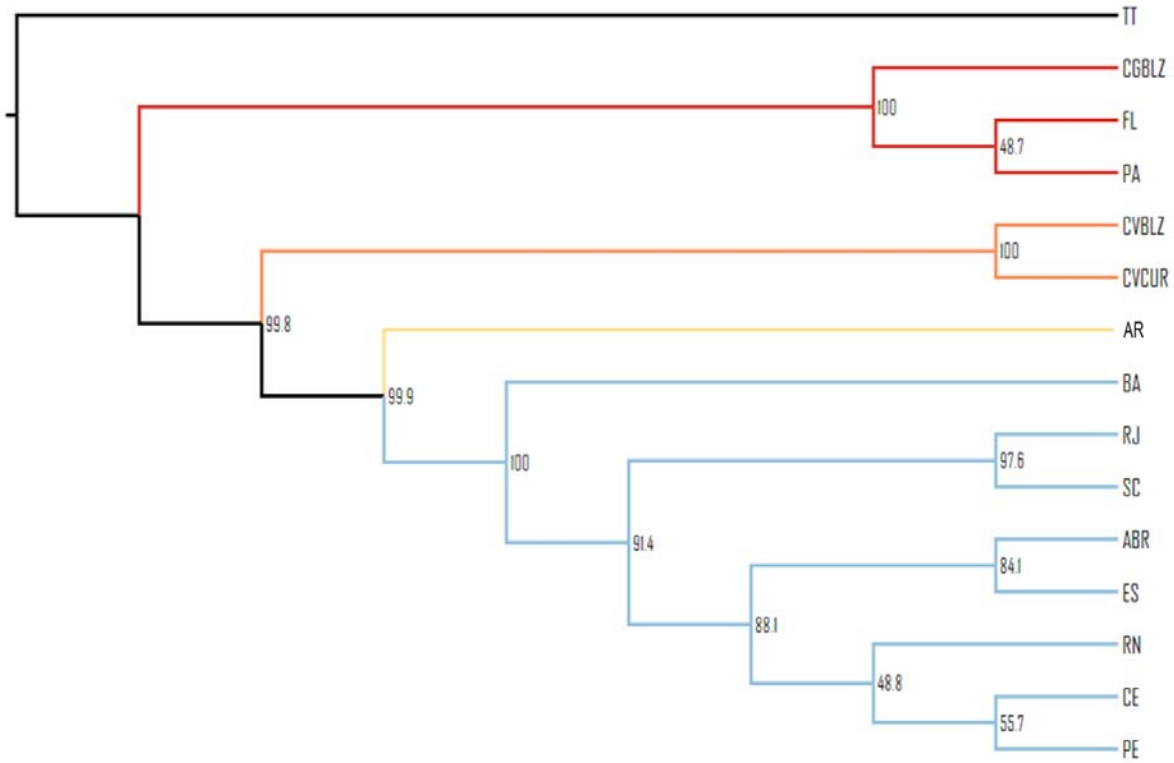

Figure 5. Multispecies coalescent tree using SNPs for all populations. Numbers at nodes represent the proportion of 1,000 bootstraps that recovered that particular node.

Within Brazil, $K = 2$ was the most likely value, which separates the two southern populations from the remainder of Brazil (Figure 7). However, there was a secondary peak in likelihood that suggested that $K = 3$ was nearly equally likely and separated Brazil into north (CE, RN and PE), central (BA, ABR and ES) and south (RJ and SC; Figure 7). With $K = 3$, these population clusters correctly correspond to my *a priori* hypothesis; no other levels of clustering were supported in Brazil (Figure 7). When evaluating barriers among Brazilian populations, both alternative hypotheses explained more variation among regions than the control AMOVA (Table 3). However, when partitioning two or three groups in Brazil, the variation explained was similar between both alternative hypotheses;

variation explained among groups for the Cabo Frio barrier was 3.9% compared to 3.6% of variation explained among groups for the Cabo Frio and SEC barriers.

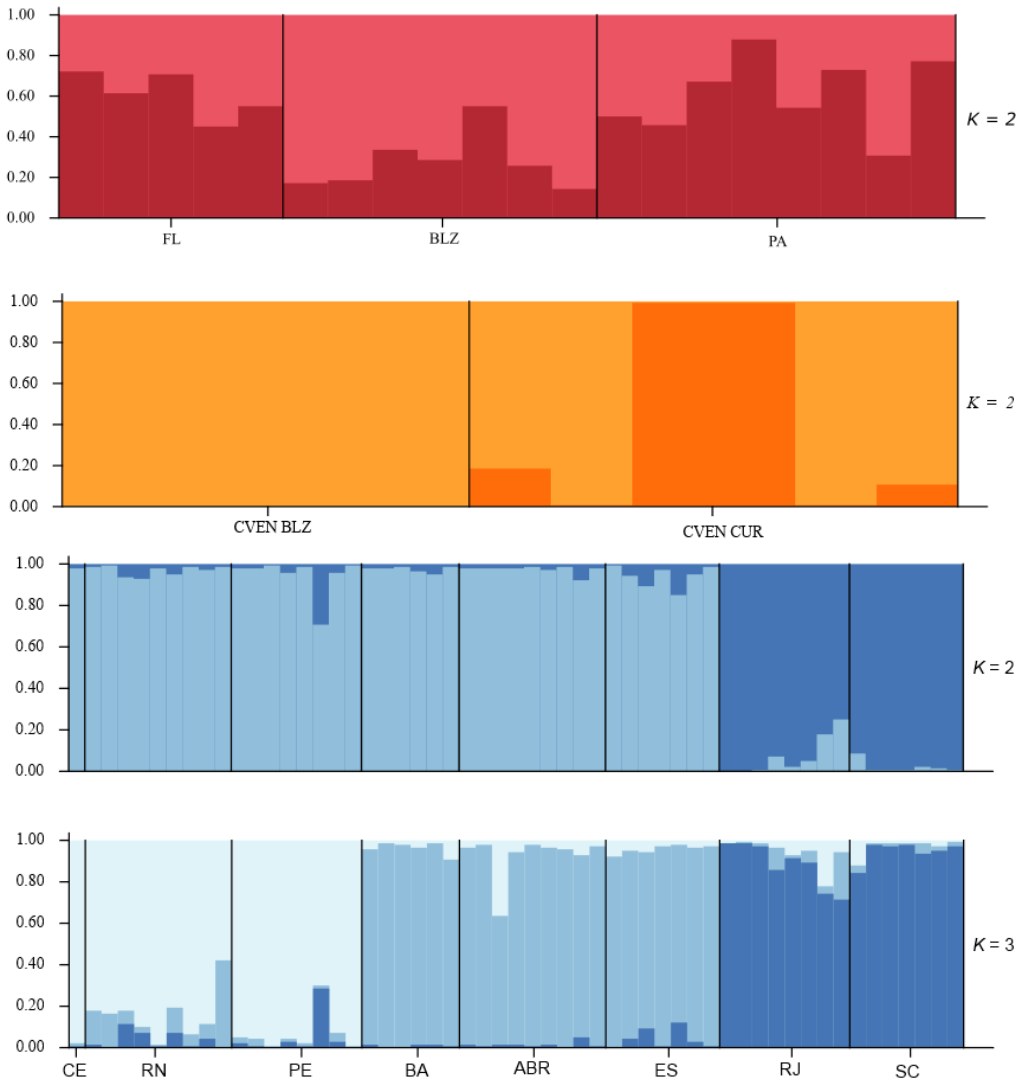

Figure 6. Results from STRUCTURE for *C. glaucofraenum*, *C. venezeuelae* and Brazil for K = 2 and Brazil for K = 3 from top to bottom.

All pairwise estimates of F_{ST} between population pairs from different clades ranged from 0.40 - 0.79 and were significantly different from zero when populations were larger than two individuals (Table 4). No pairwise estimates were significantly different from zero

between AR and any other population. However, the smallest F_{ST} comparisons were found between AR and Brazilian populations (0.28 – 0.35; excluding Ceará with only one individual; Table 4). In comparison, F_{ST} between AR and populations from *C. glaucofraenum* or *C. venezuelae* ranged from 0.40 to 0.50. Within clades, F_{ST} estimates were low (0 - 0.11), but many still showed differences significantly greater than zero (Table 4). In *C. glaucofraenum*, all three pairwise comparisons of F_{ST} were low (0.04), while the only comparison between *C. venezuelae* populations was twice as high (0.09). In Brazil, only populations across putative barriers showed a significant difference from zero, while populations in the same area were not significantly different (Table 4). In fact, when comparing pairwise F_{ST} between population pairs on the same side of a barrier and across barriers, populations across barriers showed significantly higher levels of F_{ST} ($t = 4.44$, $df = 19$, $p < 0.001$). Overall, Brazilian populations showed signs of limited dispersal based on the positive trend of IBD (Figure 8).

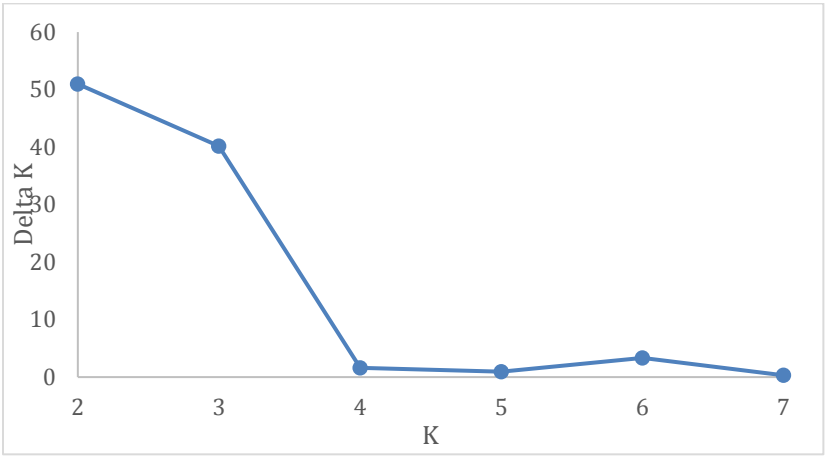

Figure 7. Plot of Δ*K* using the Evanno method. Similar likelihoods are found with $K = 2$ and $K = 3$.

When comparing levels of genetic diversity to infer recent bottleneck events, Panama (*C. glaucofraenum*) exhibited the highest levels of expected heterozygosity among all populations in any clade while Belize (*C. venezuelae*) displayed the lowest levels of heterozygosity (Table 1). Expected heterozygosity ranged from 0.090 to 0.126 and the number of effective alleles ranged from 1.131 to 1.196 across all populations (Table 1). When comparing expected heterozygosity among the three clades, there was a significant overall difference (Kruskal-Wallis $\chi^2 = 8.68$, df = 2, p-value = 0.01) with Dunn's post-hoc test indicating that both *C. venezuelae* (Z = 2.89, p > 0.01) and Brazil (Z = -2.01, p = 0.04) showed significantly lower levels of heterozygosity compared to *C. glaucofraenum*. When comparing north, central and southern Brazilian populations, heterozygosity was not different among groups (Kruskal-Wallis $\chi^2 = 1.68$, df = 2, p-value = 0.43).

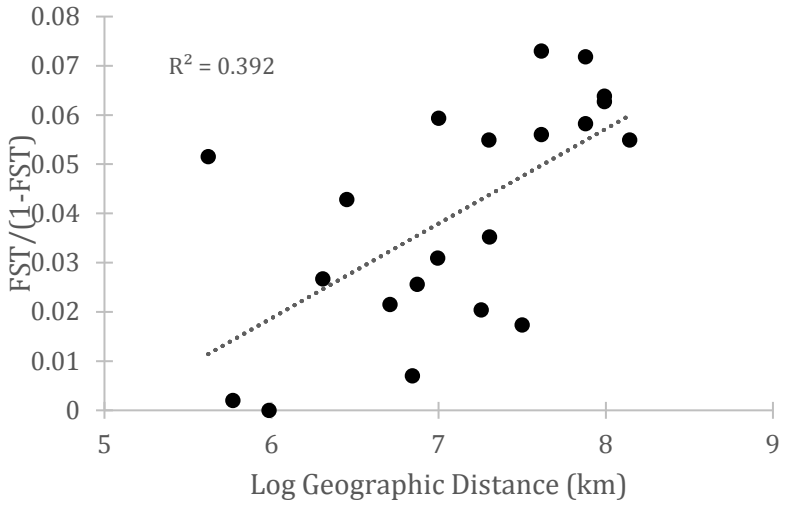

Figure 8. Isolation by distance with SNPs for all Brazilian populations except Ceará, which has only one sample. Using 10,000 permutations, distance does increase genetic divergence among populations (p = 0.0024).

CHAPTER FOUR: DISCUSSION

In this study, I was able to infer that life-history characteristics of *C. glaucofraenum* and *C. venezuelae*, such as demersal spawning and small size, would result in genetic structure throughout the western Atlantic. I expanded upon previous genetic data concerning these taxa by including several Brazilian populations and using two informative datasets to find incongruence between taxonomy and evolutionary relationships. Overall, I identified two novel clades across the Amazon barrier that are indicative of species-level genetic divergence; one clade was endemic to the Brazilian coast while the other was restricted to Atol das Rocas (AR) off the northeast coast of Brazil. In addition, minor barriers in the Caribbean show evidence of isolation between eastern and western Caribbean populations while the southern equatorial current (SEC) and Cabo Frio barriers limit gene flow among coastal Brazilian populations. Furthermore, there is indication of demographic expansion following a bottleneck event in *C. glaucofraenum* and Brazilian lineages. These results are discussed in more detail below as they relate to phylogeography of marine taxa in the western Atlantic.

Even though each monophyletic clade was strongly supported, the relationships among these lineages were discordant between the SNP and COI data. The mtDNA suggest that AR and *C. venezuelae* are more closely related, although the SNP data suggest AR and Brazil are more closely related. Although previous studies of coral reef fishes have suggested genetic connections between the Caribbean and AR due to ecologically similar environments (Lima, Freitas, Araujo, & Solé-Cava, 2005; Rocha et al., 2005), the

relationships determined using SNPs are the more likely species tree for two reasons. First, the close proximity of AR and the Brazilian coast (260 km) relative to AR and the Caribbean (>2,000 km) should allow more gene flow to occur across a short distance. Second and more importantly, sampling many genes from across the genome (as was the case with the SNP dataset) was likely to infer a more accurate species tree overall and resolve homoplasy caused by either incomplete lineage sorting that is likely to occur when analyzing only a single (mitochondrial) gene or introgression caused by occasional cross species breeding (Brito & Edwards, 2009; Edwards, Potter, Schmitt, Bragg, & Moritz, 2016).

The closely related Brazilian and AR lineages likely formed when populations were isolated from Caribbean populations due to the impact of the Amazon River outflow. The Amazon River outflow is a well-known barrier for many marine taxa and often results in speciation for low dispersal organisms like *C. glaucofraenum* and *C. venezuelae* (Bowen & Briggs, 2012; Floeter et al. 2008; Rocha, 2003). Though it is unclear when the Amazon River began to act as a barrier for marine taxa, it likely intermittently restricted gene flow for the past 9 Myr due to fluctuating sea levels that created or prevented dispersal corridors among taxa (Hoorn et al., 2017; Rocha, 2003). Based on the permeable nature of marine barriers, there was opportunity for intermittent dispersal across the Amazon, followed by periods of minimal gene flow that can result in speciation (Rocha, 2003; Floeter et al. 2008). Given that Brazil and AR diverged after *C. glaucofraenum* and *C. venezuelae* split approximately 4.21 Myr (Tornabene, Chen, & Pezold, 2013), this would put the more recent divergence of Brazil-AR from *C. venezuelae* well within the range of appearance of the Amazon barrier.

The short distance between AR and the coast should allow regular gene flow between populations to create genetic homogenization whereas more isolated habitats should have limited gene flow resulting in more unique species. Similarly isolated islands in the Atlantic exhibit endemism at half the rate as found on AR (Floeter et al. 2008). This begs the question, why is AR so differentiated from the mainland? Some studies have suggested that the level of endemism is caused by ecological differences between inshore coastal populations and offshore oceanic populations (Rocha, 2003; Rocha et al., 2005). Given that these fish exhibit a pattern of isolation-by-distance, it may be a combination of geographic distance and ecological differences that result in genetic isolation of AR.

In addition to incongruence between taxonomy and evolutionary relationships, the distribution of *C. glaucofraenum* currently extends throughout the western Atlantic to southern Brazil, where I have identified two unique clades (Brazil and AR) that correspond to species-level divergence across the Amazon barrier. In fact, divergence among all clades detected here (Table 2) was akin to species-level divergences found between other species of *Coryphopterus* (*C. hyalinus—C. personatus* = 7.16%; Baldwin, Weigt, Smith, & Mounts, 2009) and similar to the average distance between 207 other congeneric fishes (9.93%; Ward, Zemlak, Innes, Last, & Hebert, 2005). However, species delimitation should not exclusively use genetic data to define new species, so I suggest that other classes of data be incorporated here such as morphological, behavioral and ecological data (Sukumaran & Knowles, 2017).

In contrast to the species that have likely been isolated due to the presence of barriers, it is interesting that *C. glaucofraenum* and *C. venezuelae* exhibit similar levels of

genetic variation without geographic isolation. Both mate choice and ecological niche partitioning are valid explanations for the sympatric relationship between *C. glaucofraenum* and *C. venezuelae*. As fish are visually oriented, mate choice through sexual selection can help drive ecological speciation between closely related taxa (van Doorn, Edelaar, & Weissing, 2009). For instance, it was speculated that differences in vocalization may contribute to mate recognition in grunts (*Haemulon* spp.; Rocha, Rocha, Robertson, & Bowen, 2008). However, many species of *Coryphopterus* contain subtle morphological differences (Baldwin et al., 2009) and no other evidence for mate choice exists in these taxa. Therefore, sexual selection may not be driving speciation here. Alternatively, there are four ecological features that may have contributed to speciation throughout the genus due to niche partitioning. First, two of the species (*C. personatus* and *C. hyalinus*) form aggregations that hover in the water column as opposed to the other ten species which are benthic (Tornabene et al., 2013). Second, there is significant size variation with nearly half of the species in the genus (*C. tortugae, C. glaucofraenum, C. venezuelae, C. dicrus,* and *C. eilodon*) approximately twice the size of the other half (Baldwin & Robertson, 2015). Third, depth has been found to drive speciation in deep water fishes (Gaither et al., 2016) and many of these gobies vary in their maximum depth limits (Baldwin & Robertson, 2015). Fourth, while no evidence currently exists for additional partitioning of niches based on diet composition of these generalist invertivores, trophic level can play a role in divergence among closely related species within a community (Cloyed & Eason, 2017; Ferreira, Floeter, Gasparini, Ferreira, & Joyeux, 2004). The evidence for ecological speciation in

marine taxa is growing and could certainly be playing a role in diversification of fishes in the genus *Coryphopterus* (B. W. Bowen, Rocha, Toonen, & Karl, 2013; Rocha et al., 2005).

In addition to *Coryphopterus* being a diverse genus, the Caribbean is a diverse area for marine species. There are two synergistic hypotheses to explain why marine biodiversity is high in the Caribbean: 1) the Caribbean serves as a center of origin for marine speciation in the western Atlantic (Floeter et al., 2008) and 2) the Caribbean serves as a center of accumulation from nearby areas (Bowen et al. 2013; Rocha et al. 2008). With a majority of *Coryphopterus* species occurring in the Caribbean and two unique clades discovered outside the Caribbean, it is likely that dispersal occurred out of the Caribbean to Brazil in a migration event. Indeed, the fact that *C. venezuelae* can survive at a depth of 69 m below sea level suggests a possible mechanism for dispersing beyond the Amazon River Barrier. This depth is well below the water level impacted by lowered salinity from the Amazon River outflow, which typically extends to depths of 50 m (Baldwin and Robertson, 2015; Ffield, 2007).

Within clade analyses showed distinct patterns of barriers impacting connectivity throughout the species investigated in this study. The Mona Passage is typically designated as the boundary between the east and west Caribbean, although the precise location of the barrier varies (Baums, Miller, & Hellberg, 2005; DeBiasse et al., 2016; Foster et al., 2012; Taylor & Hellberg, 2003, 2006). For instance, populations of coral on opposite sides of the Mona Passage clustered together and still maintained a general east-west separation (Foster et al. 2012). Yet other studies found the Bahamas and Lesser Antilles to be genetically similar across the Mona Passage while still observing an east-west divide

(DeBiasse et al., 2016). With regard to the present study, I found evidence for a similar division among populations of *C. glaucofraenum* in the Caribbean.

One striking result of this study was the clear demarcation of the Venezuelan population as distinct from other Caribbean populations for both CGL and CVZ. Other studies have not observed such fine scale isolation across the Venezuelan coast (Betancur-R, Acero, Duque-Caro, Santos, & Knapp, 2010). Here, I found both species showed high levels of differentiation between Venezuela and the remainder of the Caribbean including a nearby population in Curacao (*C. venezuelae*). The combination of the thin continental shelf near Venezuela and observations of larvae that are transported offshore in the presence of strong oceanic currents may lead to low connectivity in both species near Venezuela (D'Agostini et al., 2015; White et al., 2010). Similarly, local currents have caused nearby populations from Belize or Honduras to show high levels of genetic differentiation over short distances (Foster et al., 2012; Jackson et al., 2014). Variability in both magnitude and location of oceanographic currents could have contributed to the relative strength and permeability of barriers, potentially contributing to isolation of Venezuela relative to other Caribbean populations.

In accordance with my prediction, two barriers were found in Brazil that genetically divide north, central and southern Brazil. The weaker of the two barriers separates northern from central Brazil and could be caused by two potential mechanisms. First, the São Francisco River outflow occurs in the same vicinity as the genetic break and has recently been referenced as a possible genetic barrier for *Millepora* fire corals and *Symbiodinium* dinoflagellates (de Souza et al., 2017; Picciani et al., 2016). Second, the

genetic boundary found between northern and central Brazil divides populations between 8°-13°S and could be due to current bifurcations. Although previous studies have primarily focused on the central southern equatorial current (cSEC) (Wieman et al., 2014), seasonal variation in both the cSEC (4-8°S) and southern SEC (sSEC; 8°-13°S) reach the genetic divide found between northern and central Brazil and could combine to cause this separation (Peterson & Stramma, 1991; Rodrigues, Rothstein, & Wimbush, 2007). Furthermore, seasonal variation in currents may help explain the weak nature of the north-central barrier. For instance, D'Agostini et al. (2015) modeled larval dispersal seasonally and found that populations in central Brazil dispersed far north in April while moving south in July due to the sSEC.

The more prominent barrier is found near Cabo Frio where the two southern populations (Rio de Janeiro and Santa Catarina) were clearly differentiated from the remaining Brazilian populations. Cabo Frio serves as the southern distribution limit for some taxa (Spalding et al., 2007) and has been known to cause differentiation in many crustaceans, but few fishes (Boschi, 2000; Fernandes et al., 2012; Maggioni et al., 2003; Santos et al., 2006). Two possible reasons for differentiation across the Cabo Frio barrier are ecological differences across the barrier or currents that prevent larval dispersal. First, the cold water and nutrient upwelling system represents an ecological transition away from warm water and live coral reefs to cool water and rocky substrate (Ferreira et al., 2004; Santos et al., 2006). Second, ocean currents may physically restrict gene flow between central and southern populations. Consistent with this mechanism, hydrodynamic modeling demonstrates the tendency for the Brazil Current to push pelagic larvae off the

continental shelf preventing larval settlement rather than following the coastline and maintaining connectivity between central and southern Brazil (D'Agostini et al. 2015).

In addition to the demographic expansion expected and found in Brazil, I also found evidence of expansion in the Caribbean. Patterns of expansion have been found in several other taxa throughout the western Atlantic. For example, multiple lines of evidence were used to show expansion of populations along the Brazilian coast (Santos, Hrbek, Farias, Schneider, & Sampaio, 2006). Expansion in the western Atlantic has been attributed to warming climate since the late Pleistocene glaciation, around 120,000 years ago (Bowen, Bass, Muss, Carlin, & Robertson, 2006). During the Pleistocene, lower sea levels and cooler temperatures may have reduced habitat availability causing populations to contract during this period (Bellwood & Wainwright, 2002). Thus, rising sea levels and warmer temperatures may have allowed habitat expansion followed by population expansion (Bowen, Bass, Muss, Carlin, & Robertson, 2006; Rodríguez-Rey, Filho, Araújo, & Solé-cava, 2017). Other studies on marine taxa have suggested the same mechanism to explain the evidence for expansion within a similar timeframe (Cunha et al., 2014; Jackson, Munguia-Vega, Beldade, Erisman, & Bernardi, 2015; Santos et al., 2006).

Overall, this study has demonstrated how genetic connectivity was impacted by permeable marine barriers throughout the western Atlantic. The Amazon River outflow has isolated Brazilian from Caribbean lineages while the offshore Brazilian archipelago of AR has also diverged from the coastal lineage. Furthermore, both COI and SNP datasets provided important information with regard to defining barriers to gene flow within regions. The mtDNA dataset provided widespread sampling throughout the range of both

Caribbean species which helped detect the east-west Caribbean barrier, whereas the SNP dataset provided in-depth information concerning the SEC and Cabo Frio barriers in Brazil that were undetectable using a single coarse marker. Lastly, evidence for demographic expansion in *C. glaucofraenum* and Brazil was found in addition to lower levels of genetic diversity in *C. venezuelae* and Brazil which indicate a potential genetic bottleneck followed by recent expansion. Overall, this study highlights how ecological barriers impact connectivity in marine taxa across the western Atlantic.

REFERENCES

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat Rev Genet*, *advance on*(2), 81–92. doi:10.1038/nrg.2015.28

Baldwin, C. C., & Robertson, D. R. (2015). A new, mesophotic *Coryphopterus* goby (Teleostei, Gobiidae) from the southern Caribbean, with comments on relationships and depth distributions within the genus. *ZooKeys*, *513*, 123–142. doi:10.3897/zookeys.513.9998

Baldwin, C. C., Weigt, L. a, Smith, D. G., & Mounts, J. H. (2009). Reconciling Genetic Lineages with Species in Western Atlantic Coryphopterus (Teleostei: Gobiidae). *Smithsonian Contributions to the Marine Sciences*, (May 2008), 111–138.

Baums, I. B., Miller, M. W., & Hellberg, M. E. (2005). Regionally isolated populations of an imperiled Caribbean coral, *Acropora palmata*. *Molecular Ecology*, *14*(5), 1377–1390. doi:10.1111/j.1365-294X.2005.02489.x

Bellwood, D. R., & Wainwright, P. C. (2002). The history and biogeography of fishes on coral reefs. *Coral Reef Fishes: Dynamics and diversity in a complex ecosystem*.

Betancur-R, R., Acero, A., Duque-Caro, H., Santos, S. R., & Knapp, M. (2010). Phylogenetic and Morphologic Analyses of a Coastal Fish Reveals a Marine Biogeographic Break of Terrestrial Origin in the Southern Caribbean. *PLoS ONE*, *5*(7). doi:10.1371/journal.pone.0011566

Blaxter, J. H. S. (2010). Fish Reproduction: *Encyclopedia of Ocean Sciences* (pp. 9–37). doi:10.1016/B978-012374473-9.00025-4

Böhlke, J. E., & Robins, C. R. (1960). A Revision of the Gobioid Fish Genus *Coryphopterus*. *Proceedings of the Academy of Natural Sciences of Philadelphia*, *112*, 103–128. Retrieved from http://www.jstor.org/stable/4064584

Boschi, E. E. (2000). Species of decapod crustaceans and their distribution in the american marine zoogeographic provinces. *Revista de Investigación Y Desarrollo Pesquero*, *13*, 1–136.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4). doi:10.1371/journal.pcbi.1003537

Bowen, B., Bass, A. L., Muss, A., Carlin, J., & Robertson, D. R. (2006). Phylogeography of two Atlantic squirrelfishes (family Holocentridae): Exploring links between pelagic larval duration and population connectivity. *Marine Biology*, *149*(4), 899–913. doi:10.1007/s00227-006-0252-1

Bowen, B. W., Rocha, L. A., Toonen, R. J., & Karl, S. A. (2013). The origins of tropical marine biodiversity. *Trends in Ecology and Evolution*. doi:10.1016/j.tree.2013.01.018

Bradbury, I. R., Laurel, B., Snelgrove, P. V. R., Bentzen, P., & Campana, S. E. (2008). Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. *Proceedings. Biological Sciences / The Royal Society*, *275*(1644), 1803–1809. doi:10.1098/rspb.2008.0216

Briggs, J. C., & Bowen, B. W. (2012). A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography*, *39*(1), 12–30. doi:10.1111/j.1365-2699.2011.02613.x

Brito, P. H., & Edwards, S. V. (2009). Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, (135), 439–455. doi:10.1007/s10709-008-9293-3

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*(23), 3317–3324. doi:10.1093/bioinformatics/btu530

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*(10), 1657–1659. doi:10.1046/j.1365-294X.2000.01020.x

Cloyed, C. S., & Eason, P. K. (2017). Niche partitioning and the role of intraspecific niche variation in structuring a guild of generalist anurans. *Royal Society of Open Science*, *4*(170060).

Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, *144*(4), 2001–2014. doi:Article

Cowen, R. K., Paris, C. B., & Srinivasan, A. (2006). Scaling of Connectivity in Marine Populations. *Science*, *311*(5760), 522–527. doi:10.1126/science.1122039

Cunha, I., Souza, A., & Dias, E. (2014). Genetic multipartitions based on D-loop sequences and chromosomal patterns in Brown chromis, *Chromis multilineata* (Pomacentridae), in the Western Atlantic. *BioMed Research International*, *2014*, 1–11. doi:http://dx.doi.org/10.1155/2014/254698

D'Agostini, A., Gherardi, D. F. M., & Pezzi, L. P. (2015). Connectivity of Marine Protected Areas and Its Relation with Total Kinetic Energy. *Plos One*, *10*(10), e0139601. doi:10.1371/journal.pone.0139601

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. doi:10.1093/bioinformatics/btr330

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772–772. doi:10.1038/nmeth.2109

de Souza, J. N., Nunes, F. L. D., Zilberberg, C., Sanchez, J. A., Migotto, A. E., Hoeksema, B. W., … Lindner, A. (2017). Contrasting patterns of connectivity among endemic and widespread fire coral species (*Millepora* spp.) in the tropical Southwestern Atlantic. *Coral Reefs*, *36*(3), 701–716. doi:10.1007/s00338-017-1562-0

DeBiasse, M. B., Richards, V. P., Shivji, M. S., & Hellberg, M. E. (2016). Shared phylogeographical breaks in a Caribbean coral reef sponge and its invertebrate commensals. *Journal of Biogeography*, *43*(11), 2136–2146. doi:10.1111/jbi.12785

Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. doi:10.1007/s12686-011-9548-7

Edwards, S. V, Potter, S., Schmitt, C. J., Bragg, J. G., & Moritz, C. (2016). Reticulation, divergence, and the phylogeography–phylogenetics continuum. *Proceedings of the National Academy of Sciences*, *113*(29), 8025–8032. doi:10.1073/pnas.1601066113

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. doi:10.1111/j.1755-0998.2010.02847.x

Faust, E., André, C., Meurling, S., Kochmann, J., Christiansen, H., Jensen, L. F., … Strand, Å. (2017). Origin and route of establishment of the invasive Pacific oyster *Crassostrea gigas* in Scandinavia. *Marine Ecology Progress Series*, *575*, 95–105. doi:10.3354/meps12219

Fernandes, D., Alves, R., Barros-alves, S. D. P., & Teixeira, G. M. (2012). Mithracinae (Decapoda : Brachyura) from the Brazilian coast : Review of the geographical distribution and comments on the biogeography of the group. *THE JOURNAL OF THE BRAZILIAN CRUSTACEAN SOCIETY*, *20*(1), 51–62.

Ferreira, C. E. L., Floeter, S. R., Gasparini, J. L., Ferreira, B. P., & Joyeux, J. C. (2004). Trophic structure patterns of Brazilian reef fishes: A latitudinal comparison. *Journal of Biogeography*, *31*(7), 1093–1106. doi:10.1111/j.1365-2699.2004.01044.x

Ffield, A. (2007). Amazon and Orinoco River plumes and NBC rings: Bystanders or participants in hurricane events? *Journal of Climate*, *20*(2), 316–333. doi:10.1175/JCLI3985.1

Floeter, S. R., Rocha, L. A., Robertson, D. R., Joyeux, J. C., Smith-Vaniz, W. F., Wirtz, P., … Bernardi, G. (2008). Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, *35*(0), 22–47. doi:10.1111/j.1365-2699.2007.01790.x

Forrester, G., Harmon, L., Helyer, J., Holden, W., & Karis, R. (2010). Experimental evidence for density-dependent reproductive output in a coral reef fish. *Population Ecology*,

53(1), 155–163. doi:10.1007/s10144-010-0225-6

Foster, N. L., Paris, C. B., Kool, J. T., Baums, I. B., Stevens, J. R., Sanchez, J. A., … Mumby, P. J. (2012). Connectivity of Caribbean coral populations: Complementary insights from empirical and modelled gene flow. *Molecular Ecology*, *21*(5), 1143–1157. doi:10.1111/j.1365-294X.2012.05455.x

Gaither, M. R., Bernal, M. A., Coleman, R. R., Bowen, B. W., Jones, S. A., Simison, W. B., & Rocha, L. A. (2015). Genomic signatures of geographic isolation and natural selection in coral reef fishes. *Molecular Ecology*, *24*(7), 1543–1557. doi:10.1111/mec.13129

Gaither, M. R., Violi, B., Gray, H. W. I., Neat, F., Drazen, J. C., Grubbs, R. D., … Hoelzel, A. R. (2016). Depth as a driver of evolution in the deep sea: Insights from grenadiers (Gadiformes: Macrouridae) of the genus Coryphaenoides. *Molecular Phylogenetics and Evolution*, *104*, 73–82. doi:10.1016/j.ympev.2016.07.027

Gaylord, B., & Gaines, S. D. (2000). Temperature or Transport? Range Limits in Marine Species Mediated Solely by Flow. *The American Naturalist*, *155*(6), 769–789. doi:10.1086/303357

Gleason, L. U., & Burton, R. S. (2016). Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail *Chlorostoma funebralis*. *Molecular Ecology*, *25*(15), 3557–3573. doi:10.1111/mec.13703

Gottscho, A. D., Wood, D. A., Vandergast, A. G., Lemos-Espinal, J., Gatesy, J., & Reeder, T. W. (2017). Lineage diversification of fringe-toed lizards (Phrynosomatidae: *Uma notata* complex) in the Colorado Desert: Delimiting species in the presence of gene flow. *Molecular Phylogenetics and Evolution*, *106*, 103–117.

doi:10.1016/j.ympev.2016.09.008

Hoorn, C., Bogotá-A, G. R., Romero-Baez, M., Lammertsma, E. I., Flantua, S. G. A., Dantas, E. I., … Chemale, F. (2017). The Amazon at sea: Onset and stages of the Amazon River from a marine record in the Foz do Amazonas Basin (Brazilian Equatorial Margin), with special reference to vegetation turnover in the Plio-Pleistocene. *Global and Planetary Change*. doi:10.1016/j.gloplacha.2017.02.005

Jackson, A. M., Munguia-Vega, A., Beldade, R., Erisman, B. E., & Bernardi, G. (2015). Incorporating historical and ecological genetic data for leopard grouper (*Mycteroperca rosacea*) into marine reserve design in the Gulf of California. *Conservation Genetics*, *16*(4), 811–822. doi:10.1007/s10592-015-0702-8

Jackson, A. M., Semmens, B. X., Sadovy De Mitcheson, Y., Nemeth, R. S., Heppell, S. A., Bush, P. G., … Yue, G. H. (2014). Population Structure and Phylogeography in Nassau Grouper (Epinephelus striatus), a Mass-Aggregating Marine Fish. *PLoS ONE*, *9*(5). doi:10.1371/journal.pone.0097508

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, msw054. doi:10.1093/molbev/msw054

Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*(6), 1695–1701. doi:10.1093/molbev/mss020

Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. doi:10.1111/2041-

210X.12410

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, *50*(6), 913–925. doi:10.1080/106351501753462876

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. doi:10.1093/bioinformatics/btr509

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. doi:10.1093/bioinformatics/btp352

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451–1452. doi:10.1093/bioinformatics/btp187

Lima, D., Freitas, J. E. P., Araujo, M. E., & Solé-Cava, A. M. (2005). Genetic detection of cryptic species in the frillfin goby *Bathygobius soporator*. *Journal of Experimental Marine Biology and Ecology*, *320*(2), 211–223. doi:10.1016/j.jembe.2004.12.031

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299. doi:10.1093/bioinformatics/btr642

Luo, S.-J., Kim, J.-H., Johnson, W. E., van der Walt, J., Martenson, J., Yuhki, N., … O'Brien, S. J. (2004). Phylogeography and genetic ancestry of tigers (*Panthera tigris*). *PLoS Biology*, *2*(12), e442. doi:10.1371/journal.pbio.0020442

Machado, L. F., Damasceno, J. de S., Bertoncini, Á. A., Tosta, V. C., Farro, A. P. C., Hostim-Silva, M., & Oliveira, C. (2017). Population genetic structure and demographic history of the spadefish, *Chaetodipterus faber* (Ephippidae) from Southwestern Atlantic. *Journal of Experimental Marine Biology and Ecology*, *487*, 45–52. doi:10.1016/j.jembe.2016.11.005

Maggioni, R., Rogers, a. D., & Maclean, N. (2003). Population structure of *Litopenaeus schmitti* (Decapoda: Penaeidae) from the Brazilian coast identified using six polymorphic microsatellite loci. *Molecular Ecology*, *12*, 3213–3217. doi:10.1046/j.1365-294X.2003.01987.x

Magsino, R. M., & Juinio-Meñez, M. A. (2008). The influence of contrasting life history traits and oceanic processes on genetic structuring of rabbitfish populations *Siganus argenteus* and *Siganus fuscescens* along the eastern Philippine coasts. *Marine Biology*, *154*(3), 519–532. doi:10.1007/s00227-008-0946-7

Milá, B., Van Tassell, J. L., Calderón, J. A., Rüber, L., & Zardoya, R. (2017). Cryptic lineage divergence in marine environments: genetic differentiation at multiple spatial and temporal scales in the widespread intertidal goby *Gobiosoma bosc*. *Ecology and Evolution*, *7*(14), 5514–5523. doi:10.1002/ece3.3161

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES science gateway for inference of large phylogenetic trees. In 2010 Gateway Computing Environments Workshop, GCE 2010.

Momigliano, P., Jokinen, H., Fraimout, A., Florin, A.-B., Norkko, A., & Merilä, J. (2017). Extraordinarily rapid speciation in a marine fish. *Proceedings of the National Academy*

of Sciences, *114*(23), 201615109. doi:10.1073/pnas.1615109114

Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: An empirical study from an Amazonian plant species. *Molecular Ecology Resources*. doi:10.1111/1755-0998.12654

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, *25*(1994), 547–572. doi:10.1146/annurev.es.25.110194.002555

Peterson, R. G., & Stramma, L. (1991). Upper-level circulation in the South-Atlantic Ocean. *Progress In Oceanography*, *26*(1), 1–73. doi:10.1016/0079-6611(91)90006-8

Picciani, N., de Lossio e Seiblitz, I. G., de Paiva, P. C., e Castro, C. B., & Zilberberg, C. (2016). Geographic patterns of Symbiodinium diversity associated with the coral *Mussismilia hispida* (Cnidaria, Scleractinia) correlate with major reef regions in the Southwestern Atlantic Ocean. *Marine Biology*, *163*(11). doi:10.1007/s00227-016-3010-z

Prates, I., Xue, A. T., Brown, J. L., Alvarado-Serrano, D. F., Rodrigues, M. T., Hickerson, M. J., & Carnaval, A. C. (2016). Inferring responses to climate dynamics from historical demography in neotropical forest lizards. *Proceedings of the National Academy of Sciences*, 1–8. doi:10.1073/pnas.1601063113

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. doi:10.1111/j.1471-8286.2007.01758.x

Rambaut A (2014) FigTree v1.4.2. http://tree.bio.ed.ac.uk/soft- ware/figtree/

Riginos, C., Douglas, K. E., Jin, Y., Shanahan, D. F., & Treml, E. A. (2011). Effects of geography

and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, *34*(4), 566–575. doi:10.1111/j.1600-0587.2010.06511.x

Robins, C.R., Ray, G.C., 1986. A Field Guide to Atlantic Coast Fishes of North America. Houghton Mifflin, Boston

Rocha, L. A. (2003). Patterns of distribution and processes of speciation in Brazilian reef fishes. *Journal of Biogeography*, *30*, 1161–1171.

Rocha, L. A., Robertson, D. R., Roman, J., & Bowen, B. W. (2005). Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1563), 573–579. doi:10.1098/2004.3005

Rocha, L. A., Rocha, C. R., Robertson, D. R., & Bowen, B. W. (2008). Comparative phylogeography of Atlantic reef fishes indicates both origin and accumulation of diversity in the Caribbean. *BMC Evolutionary Biology*, *8*, 157. doi:10.1186/1471-2148-8-157

Rodrigues, R. R., Rothstein, L. M., & Wimbush, M. (2007). Seasonal Variability of the South Equatorial Current Bifurcation in the Atlantic Ocean: A Numerical Study. *Journal of Physical Oceanography*, *37*(1), 16–30. doi:10.1175/JPO2983.1

Rodríguez-Rey, G. T., Filho, A. C., Araújo, M. E. D. E., & Solé-cava, A. M. (2017). Evolutionary history of *Bathygobius* (Perciformes : Gobiidae) in the Atlantic biogeographic provinces : a new endemic species and old mitochondrial lineages. *Zoological Journal of the Linnean Society*, 1–25.

Rohland, N., & Reich, D. (2012). Cost-effective , high-throughput DNA sequencing. *Genome Research*, *22*, 939–946. doi:10.1101/gr.128124.111

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*(12), 1572–1574. doi:10.1093/bioinformatics/btg180

Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106. doi:10.1111/j.1471-8286.2007.01931.x

Roux, C., Fraïsse, C., Romiguier, J., Anciaux, Y., Galtier, N., Bierne, N., … Blum, M. (2016). Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence. *PLOS Biology*, *14*(12), 1–22. doi:10.1371/JOURNAL.PBIO.2000234

RStudio, 2013. RStudio: integrated development environment for R, Version 0.97.332. Boston, MA, Available at <http://www.rstudio.org>.

Russello, M. A., Waterhouse, M. D., Etter, P. D., & Johnson, E. A. (2015). From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ*, *3*, e1106. doi:10.7717/peerj.1106

Saenz-Agudelo, P., Dibattista, J. D., Piatek, M. J., Gaither, M. R., Harrison, H. B., Nanninga, G. B., & Berumen, M. L. (2015). Seascape genetics along environmental gradients in the Arabian Peninsula: Insights from ddRAD sequencing of anemonefishes. *Molecular Ecology*, 6241–6255. doi:10.1111/mec.13471

Santos, S., Hrbek, T., Farias, I. P., Schneider, H., & Sampaio, I. (2006). Population genetic structuring of the king weakfish, *Macrodon ancylodon* (Sciaenidae), in Atlantic coastal waters of South America: deep genetic divergence without morphological change. *Molecular Ecology*, *15*(14), 4361–73. doi:10.1111/j.1365-294X.2006.03108.x

Slatkin, M. (1985). Gene Flow in Natural Populations. *Annual Review of Ecology and Systematics*, *16*, 393–430. Retrieved from http://www.annualreviews.org/doi/pdf/10.1146/annurev.es.16.110185.002141

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Nature*, *236*(4803), 787–792. Retrieved from http://www.jstor.org/stable/1699930

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. a., Finlayson, M., … Robertson, J. (2007). Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. *BioScience*, *57*(7), 573. doi:10.1641/B570707

Spinks, P. Q., Thomson, R. C., & Shaffer, H. B. (2014). The advantages of going large: Genome-wide SNPs clarify the complex population history and systematics of the threatened western pond turtle. *Molecular Ecology*, *23*(9), 2228–2241. doi:10.1111/mec.12736

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. doi:10.1093/bioinformatics/btu033

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *2016*, 201607921. doi:10.1073/PNAS.1607921114

Swofford DL (2002) PAUP* Phylogenetic Analysis Using Parsi- mony (*and Other Methods), Version 4. Sinauer Associates, Sunderland, MA.

Tamura, K., & Nei, M. (1993). Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial DNA in Humans and Chimpanzees. *Molecular Biology*

and Evolution, *10*(3), 512–526.

Taylor, M. S., & Hellberg, M. E. (2003). Genetic Evidence for Local Retention of Pelagic Larvae in a Caribbean Reef Fish. *Science*, *299*(5603), 107–109. doi:10.1126/science.1079365

Taylor, M. S., & Hellberg, M. E. (2006). Comparative phylogeography in a genus of coral reef fishes: Biogeographic and genetic concordance in the Caribbean. *Molecular Ecology*, *15*(3), 695–707. doi:10.1111/j.1365-294X.2006.02820.x

Tornabene, L., Chen, Y. J., & Pezold, F. (2013). Gobies are deeply divided: phylogenetic evidence from nuclear DNA (Teleostei: Gobioidei: Gobiidae). *Systematics and Biodiversity*, *11*(3), 345–361. doi:Doi 10.1080/14772000.2013.818589

van Doorn, S. G., Edelaar, P., & Weissing, F. J. (2009). On the Origin of Species by Natural and Sexual Selection. *Science*, *326*(5960), 1704–1707. doi:10.1126/science.1178883

Villanova, V. L., Hughes, P. T., & Hoffman, E. A. (2017). Combining genetic structure and demographic analyses to estimate persistence in endangered Key deer (*Odocoileus virginianus clavium*). *Conservation Genetics*, (18), 1061–1076. doi:10.1007/s10592-017-0958-2

Ward, R., Zemlak, T., Innes, B., Last, P., & Hebert, P. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1847–1857. doi:10.1098/rstb.2005.1716

White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C., & Toonen, R. J. (2010). Ocean currents help explain population genetic structure. *Proc Biol Sci*, *277*(1688), 1685–1694. doi:10.1098/rspb.2009.2214

Wieman, A. C., Berendzen, P. B., Hampton, K. R., Jang, J., Hopkins, M. J., Jurgenson, J., … Thurman, C. L. (2014). A panmictic fiddler crab from the coast of Brazil? Impact of divergent ocean currents and larval dispersal potential on genetic and morphological variation in Uca maracoani. *Marine Biology*, *161*(1), 173–185. doi:10.1007/s00227-013-2327-0

Yamazaki, D., Miura, O., Ikeda, M., Kijima, A., Van Tu, D., Sasaki, T., & Chiba, S. (2017). Genetic diversification of intertidal gastropoda in an archipelago: the effects of islands, oceanic currents, and ecology. *Marine Biology*, *164*(9). doi:10.1007/s00227-017-3207-9

Young, C. R., Fujio, S., & Vrijenhoek, R. C. (2008). Directional dispersal between mid-ocean ridges: Deep-ocean circulation and gene flow in *Ridgeia piscesae*. *Molecular Ecology*, *17*(7), 1718–1731. doi:10.1111/j.1365-294X.2008.03609.x