Received: 11 September 2019 | Revised: 17 July 2020 | Accepted: 24 September 2020

DOI: 10.1111/jbi.14010

RESEARCH PAPER

Journal of Biogeography | WILEY

Going against the flow: Barriers to gene flow impact patterns of connectivity in cryptic coral reef gobies throughout the western Atlantic

Daniel R. Volk[1,2] | John D. Konvalina[1] | Sergio R. Floeter[3] | Carlos E. L. Ferreira[4] | Eric A. Hoffman[1]

[1]University of Central Florida, Orlando, Florida, USA

[2]Cleveland Metroparks, Parma, OH, USA

[3]Marine Macroecology and Biogeography Lab, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil

[4]Marine Biology Department, Universidade Federal Fluminense, Niterói, RJ, Brazil

Correspondence
Eric A. Hoffman, 4000 Central Florida Blvd., Department of Biology, University of Central Florida, Orlando, FL 32816, USA.
Email: eric.hoffman@ucf.edu

Funding information
Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina, Grant/Award Number: 6308/2011-8; Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number: 563276/2010-0

Handling Editor: Aristeidis Parmakelis

Abstract

Aim: Complex oceanographic features have historically caused difficulty in understanding gene flow in marine taxa. Here, we evaluate the impact of potential phylogeographic barriers to gene flow and assess demography and evolutionary history of a coral reef goby species complex. Specifically, we test how the Amazon River outflow and ocean currents impact gene flow.

Location: Western Atlantic.

Taxon: The bridled goby (*Coryphopterus glaucofraenum*) and sand-canyon goby (*C. venezuelae*) species complex.

Methods: We used mitochondrial DNA and 2,401 genomic SNPs to investigate evolutionary history and test hypotheses of how major barriers impact species-level differentiation. We used clustering algorithms and pairwise F_{ST} to assess population differentiation caused by minor barriers within and among regions. Finally, we tested alternate hypotheses of demographic history via coalescent simulations to determine the most plausible spread across the Western Atlantic.

Results: We found two unique clades of *C. glaucofraenum* along the Brazilian coast and the oceanic island Atol das Rocas (AR) that are more closely related to *C. venzuelae*. Further genetic structure within the Caribbean and separately along the Brazilian coast led to at least two distinct populations in each Province. Coalescent simulations indicated that an ancestral population of *C. venezuelae* split from *C. glaucofraenum* in the Caribbean, dispersed to Brazil, then spread to AR.

Main conclusions: Species-level genetic differentiation has resulted from the Amazon River outflow and isolation of AR. Population differentiation within the Caribbean matched previous studies indicating an east-west pattern of divergence. Brazilian population differentiation was impacted by the cold-water upwelling filter at Cabo Frio. Overall, this research highlights how barriers to gene flow impact speciation and genetic structure within western Atlantic gobies and provides insight into the role oceanographic features have in the speciation process of fishes.

KEYWORDS

Brazil, Caribbean, *Coryphopterus*, cryptic species, marine speciation, SNP

 © 2020 John Wiley & Sons Ltd

1 | INTRODUCTION

Gene flow is an important biological process that can lead to genetic homogenization of populations and prevent taxon divergence (Tigano & Friesen, 2016). Although marine systems exhibit few obvious biogeographic barriers, at least four major factors are known to impact among-population isolation throughout the oceanic realm. First, life-history characteristics such as spawning modes influence overall dispersal potential such that demersal spawners typically exhibit greater genetic structure than pelagic spawners (e.g. Floeter et al., 2008). Second, currents can also act as barriers that prevent larval dispersal (Gaylord & Gaines, 2000), and lead to genetic isolation (e.g. Huyghe & Kochzius, 2018; Santos et al., 2006). Third, freshwater and sediment outflow from rivers can create impassable biogeographic barriers to some taxa (Rocha, 2003). Fourth, genetic isolation can result among populations living in geographically isolated habitats, such as oceanic islands (Dias et al., 2019).

The western Atlantic is characterized by several biogeographic barriers and is thus divided into Caribbean and Brazilian biogeographic provinces (Floeter et al., 2008). At the largest scale, marked differences in species composition between the Caribbean and Brazilian provinces have been attributed to the Amazon River outflow which acts as an ecological barrier to gene flow (Briggs & Bowen, 2012; Rocha, 2003). At finer spatial scales, genetic patterns show an east-west division within the Caribbean and isolation of the Bahamas that are likely due to ocean currents (DeBiasse et al., 2016; Foster et al., 2012; Jackson et al., 2014; Taylor & Hellberg, 2006). Within the Brazilian province, genetic structure among populations of decapods, photosynthetic dinoflagellates and fish has been attributed to the Southern Equatorial Current (SEC) and Cabo Frio barriers resulting in northern, central and southern clusters (Boschi, 2000; Picciani et al., 2016; Santos et al., 2006).

Species with lower dispersal potential (i.e. demersal spawners) are model species to evaluate the impact of biogeographic barriers throughout the western Atlantic. Two sister species, the bridled goby (*Coryphopterus glaucofraenum* Gill 1863) and the sand-canyon goby (*C. venezuelae* Cervigón 1966), are small (<55 mm), benthic fishes that lay demersal eggs in nests on sandy patches near coral reefs which they defend (Forrester et al., 2010). Both *C. glaucofraenum* and *C. venezuelae* occur throughout the Caribbean while *C. glaucofraenum* extends to southern Brazil (Robins & Ray 1986). Early studies of *C. glaucofraenum* described subtle morphological variation in bone and ray counts as well as coloration among western Atlantic populations (Böhlke & Robins, 1960). Subsequent studies elevated multiple subspecies of *C. glaucofraenum* to full species designation, including *C. venezuelae*, based on genetic and morphological characteristics (Baldwin et al., 2009). However, morphology is often similar between species within *Coryphopterus*, particularly larval morphology, which commonly leads to misidentification. Because most gobies are demersal spawners and often demonstrate significant genetic structure across their geographic range (Milá et al., 2017), it is likely that whole-range genetic sampling can facilitate taxonomic differentiation between closely related, sympatric species.

Here, we sampled across a broad geographic scale and employed mtDNA sequence data and thousands of single nucleotide polymorphisms (SNPs) to test if species- or population-level differences were present within *C. glaucofraenum* and *C. venezuelae*. We hypothesized that the Amazon River outflow and the isolation of the Brazilian oceanic reefs would create unique genetic clades indicative of species-level genetic divergence among populations on either side of these major barriers. Second, we hypothesized that minor barriers in both the Caribbean and Brazilian provinces would promote population structure such that populations within the eastern or western Caribbean would be more genetically similar to each other than populations compared across the Caribbean. Similarly, we hypothesized that coastal Brazilian populations would be separated into three genetic clusters based on the SEC and Cabo Frio barriers. Finally, we used coalescent simulations to test how dispersal occurred across these barriers.

2 | METHODS

2.1 | COI sequence analyses

We collected 112 individuals of *C. glaucofraenum* across 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 placed in 95% ethanol and frozen for long-term storage. We extracted genomic DNA using a Serapure bead protocol (Rohland & Reich, 2012) and amplified a 690 bp alignment of the COI gene with FishF1 and FishR1 primers using slight modifications (i.e. annealing temp was 35 s at 53°C) from Ward et al. (2005). PCR products were sent to Eurofins Genomics for sequencing. We verified chromatographs by eye using SEQUENCHER version 5.1 (Gene Codes, Ann Arbor, MI, USA). Sequences were then trimmed and aligned with GenBank samples using MEGA version 7 (Kumar et al., 2016) followed by file formatting for each analysis using PGDSPIDER (Lischer & Excoffier, 2012).

In order to determine the evolutionary relationships among lineages, we performed a Bayesian phylogenetic analysis using BEAST2 (Bouckaert et al., 2014) with the HKY + G model of evolution as determined in PartitionFinder (Lanfear et al., 2012). We 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 version 1.6 (Bouckaert et al., 2014) to ensure effective sample sizes (ESS) were ≥200 for each parameter. LOGCOMBINER version 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 9,000 tree states. Using this combined file, we used TREEANNOTATOR to create a 50% majority-rule consensus tree which was viewed in FIGTREE version 1.4.2 (Rambaut, 2016). We considered Bayesian posterior support values >0.95 to be indicative of highly supported nodes. We also created a TCS (Clement et al., 2000) haplotype network in POPART (Leigh & Bryant, 2015) to visualize the distribution of haplotypes among clades and populations.

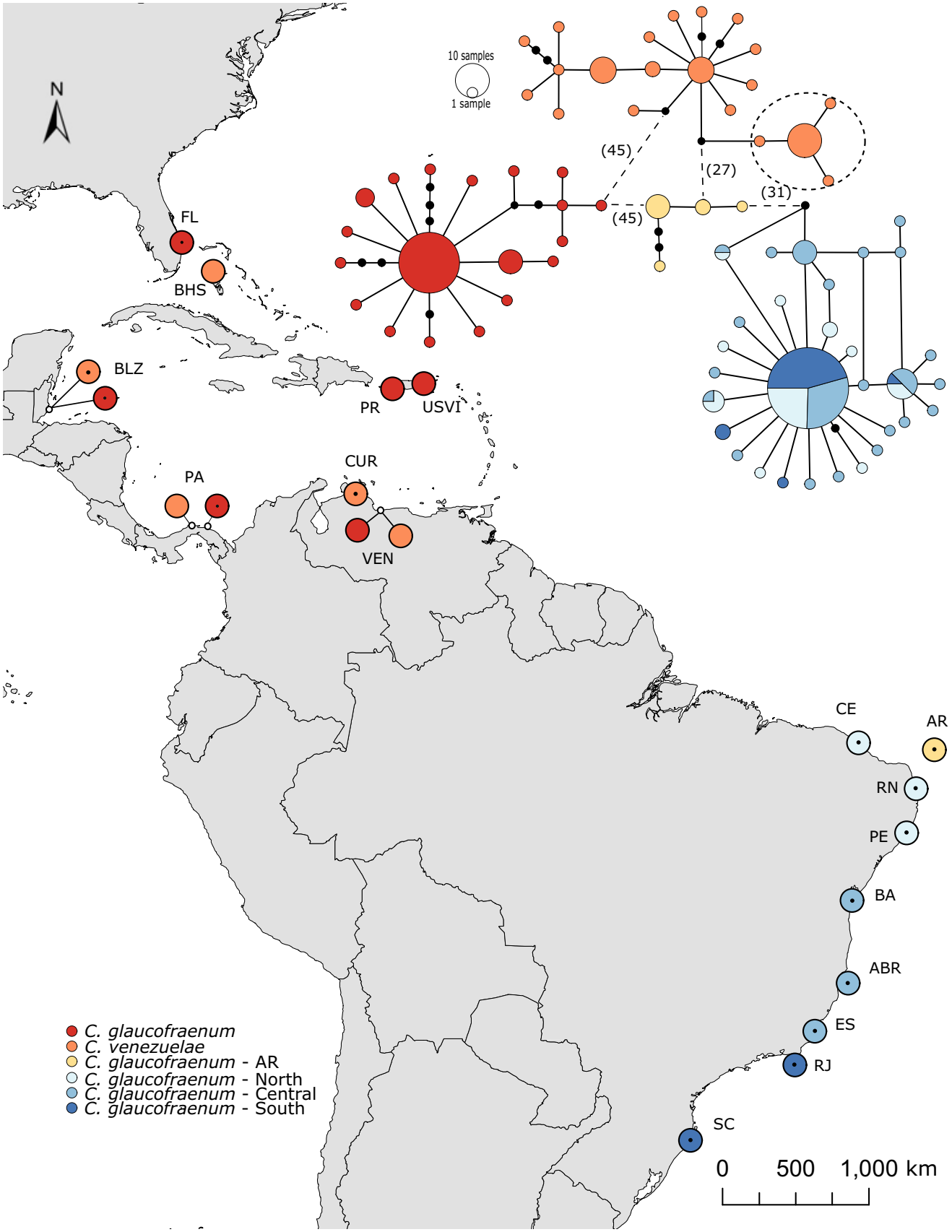

FIGURE 1 Western Atlantic map of populations of *C. glaucofraenum* and *C. venezuelae* used in this study. All sampling points include COI data while circles with dots indicate that SNPs were also used. Haplotype network represents COI data with shading that illustrates populations. Dashed lines with numbers indicate the number of inferred mutations between lineages and the three unique haplotypes circled within *C. venezuelae* show individuals collected from Venezuela

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 and summarized with individuals from all populations. 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 Overall (*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	–	–	–	–	–	–
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})	–	–	–	–
Curaçao (CVEN CUR)	10	7	0.911 (0.077)	3.0×10^{-3} (0.62×10^{-3})	6	1.136	0.075	0.091
Caribbean Overall (*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	–	–	–	–	–	–
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 Overall	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	–	–	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

2.2 | COI population genetic analyses

We first summarized basic genetic diversity estimates. In Arlequin version 3.5.2 (Excoffier & Lisher, 2010), we estimated pairwise distance (F_{ST}) with 20,000 permutations and the Tamura and Nei (1993) substitution model (HKY + G substitution model was not available in Arlequin). Within each clade, we expected populations within an area to be more similar to each other than populations across a barrier. Thus, we compared pairwise distance in *C. glaucofraenum* between populations within the east (USVI, PR, VEN, CUR) and west (FL, BLZ, PA) Caribbean to pairwise distance between east-west population pairs using a student's *t* test. Sparse population sampling prohibited a similar analysis for *C. venezuelae*.

To evaluate whether barriers impact population connectivity in Brazil, we also tested whether *C. glaucofraenum* populations within northern, central or southern Brazil were more similar to each other than population pairs across these regions using a Wilcoxon rank-sum test.

2.3 | SNP generation and filtering

A reduced-sample SNP dataset was generated using 103 total 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 et al. (2015). The nextRAD libraries were sequenced on an Illumina HiSeq 4,000 with one lane of single-end 150 bp reads (University of Oregon). After de novo assembly and initial filtering, there were 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 (Table S1). Detailed information on the standard SNPsaurus SNP generation and filtering can be found in Supplemental Document 1. To identify loci under selection, we used the program Bayescan v.2.1 (Foll & Gaggiotti, 2008) using default parameters for a neutral model: 100,000 iterations and 10 prior odds. The maximum allowable false discovery rate (FDR) was set to 0.05. We ran Bayescan within each putative species (i.e. Brazilian C. *glaucofraenum* (with southern Brazil versus central and northern Brazil); *C. venezualae* and Caribbean *C. glaucofraenum*; see Results). All outlier loci were removed from downstream analyses.

2.4 | SNP phylogenetic analyses

To estimate evolutionary relationships among species, we first utilized a Bayesian approach in MrBayes version 3.2.6 (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 version 2.1.10 (Darriba et al., 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). Additionally, we performed a maximum likelihood analysis in RAxML (Stamatakis, 2014) using concatenated loci and a correction bias due to using all variable sites (Lewis, 2001). Here, only two samples from AR with large amounts of missing data were constrained while all other taxa were not. Both trees were visualized and modified in FigTree.

2.5 | SNP population genetic analyses

To obtain a comprehensive understanding of the genetic diversity present, we first summarized basic genetic diversity metrics for populations for the SNP dataset. To evaluate population structure within clades identified in the phylogenetic analysis, a Bayesian clustering analysis was performed within each clade and without population location priors in Structure (Pritchard et al., 2000). Here, we performed 10 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 et al., 2005) was used in StructureHarvester (Earl & VonHoldt, 2012) to determine the most likely value for K. After initial runs were complete, we checked for substructure by rerunning Structure within genetic clusters using the same parameters. Because Structure results were ambiguous for Brazil *C. glaucofraenum* populations (see Results), we further assessed population structure among these populations via a Discriminant Analysis of Principal Components (DAPC) using the R package 'adegenet' v.2.1.2 (Jombart et al., 2010). Unlike Structure, DAPC is free of model-based assumptions and aims to maximize variance between groups while minimizing variance within groups. The data were transformed into principal components (PCs) in order to decrease the number of variables and reduce computing time. After choosing the minimum number of PCs necessary to describe the variation in the data, we selected the optimal number of discriminant functions (DFs) to describe the number of clusters.

To determine population differentiation, we estimated pairwise F_{ST} among populations using the pairwise distance approach in Arlequin following the same approach as with COI data. We then compared levels of F_{ST} between population pairs from the same area to estimates of F_{ST} between population pairs from different regions along the Brazilian coast using a student's t test. Specifically, we tested to see if populations within northern, central or southern Brazil were more similar to each other than population pairs from different areas of Brazil. Additionally, Brazilian populations were tested for isolation by distance (IBD) to see if populations were dispersal limited using GenePop (Rousset, 2008). Other regions had too few populations to make similar comparisons.

Lastly, we tested four of the most plausible alternative hypotheses of the demographic history of these lineages (Figure 2). Model 1 posits a common ancestral population that gave rise to all three lineages simultaneously. Model 2 illustrates a dispersal event from the Caribbean to a collective Brazil/AR ancestral population which later diverged. Model 3 hypothesizes a sequence of dispersal events from the Caribbean to the Brazilian coast, and from Brazil to AR. Model 4 allows us to test an alternative sequence of dispersal events such that dispersal occurred from the Caribbean to AR first, then to the Brazilian coast. Overall, the first model serves as a null hypothesis and the other three models are potential scenarios given what we know about the Caribbean serving as a hotbed of speciation spreading diversity to other Atlantic regions (Lima et al., 2005; Rocha et al., 2005). To evaluate between these models, we used coalescent simulations in the program fastsimcoal2 version 2.3.0.3 (Excoffier et al., 2013). We first converted the variant call format (VCF) file containing all the SNPs to fastsimcoal2 format following Liu et al. (2018). This gave us the observed site frequency spectrum (SFS). Because we did not have an outgroup, we used a folded SFS showing minor allele frequencies. With fastsimcoal2 we ran 10 independent runs of each model, with each run containing 5,000 simulations and 20 expectation-conditional maximization (ECM) cycles. Of the 10 independent runs, we chose the run with the least distance between observed and expected maximum likelihood to be used in AIC model selection. AIC compared the four models and selected the model with the highest likelihood of obtaining the observed site frequency spectrum.

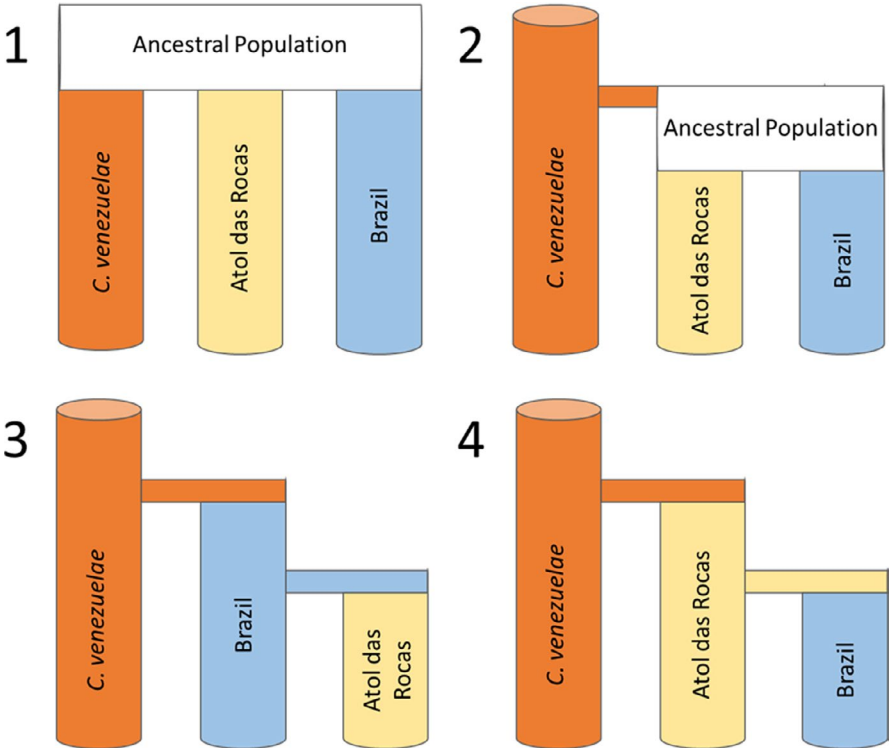

FIGURE 2 Alternative demographic models of *C. glaucofraenum* and *C. venezuelae* tested using FASTSIMCOAL2. Model 1 illustrates a vicariant event splitting all three lineages simultaneously. Model 2 shows dispersal of *C. venezuelae* from the Caribbean followed by a vicariant event which splits AR and Brazil. Model 3 illustrates dispersal to Brazil first, then from Brazil to AR. Model 4 tests the opposite dispersal, i.e. from the Caribbean to AR, then AR to Brazil

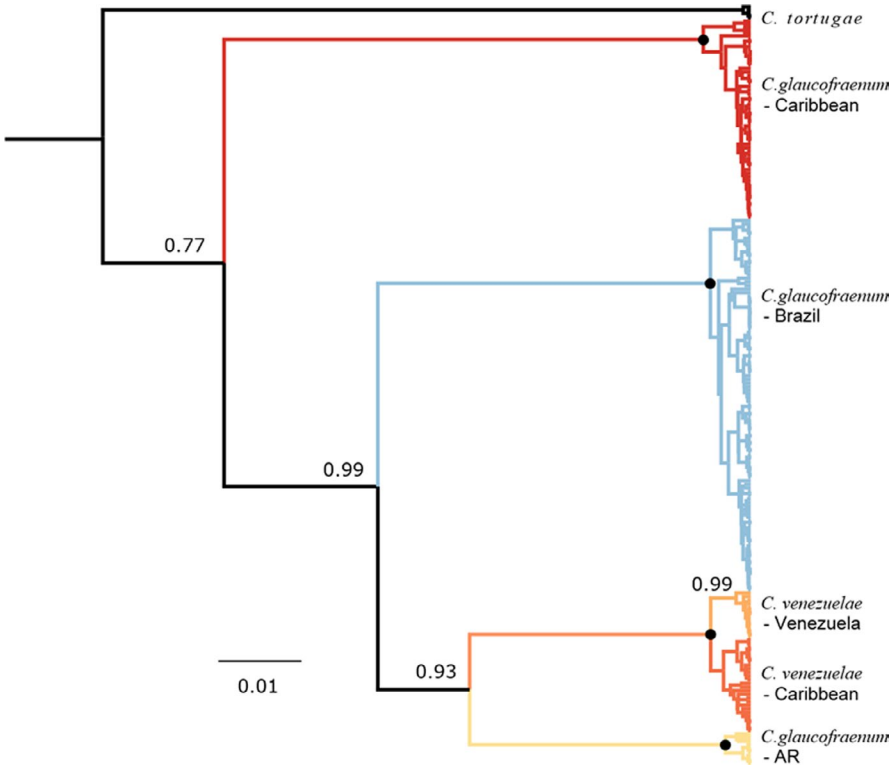

FIGURE 3 Bayesian phylogeny of *C. glaucofraenum* and *C. venezuelae* using COI from BEAST2 with clades coloured to represent collection sites. Black circles represent a posterior value of 1 and any posterior support values > 0.75 are shown at nodes. There is strong support (0.99) to show that AR *C. glaucofraenum* and Brazil *C. glaucofraenum* both share a more recent common ancestor with *C. venezuelae* than either do with Caribbean *C. glaucofraenum*

3 | RESULTS

3.1 | COI sequence analyses

We found four highly supported monophyletic clades with all four highly divergent from one another (Figure 3). While the basal node showed poor support, most other nodes exhibited high support (> 0.95 posterior support). Even though Brazil and AR (both "*C. glaucofraenum*") are close in proximity, these results show strong support to suggest that *C. glaucofraenum* from Brazil and AR are more closely related to *C. venezuelae* than either clade is to Caribbean *C. glaucofraenum*. Moreover, per cent sequence divergence among the four

TABLE 2 Pairwise F_{ST} estimates for COI (lower) and SNP data (upper) among all locations. Values in bold are significant at $p < 0.05$ for COI and $p < 0.01$ for SNPs 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		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	0		**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	0	0	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	0.01	0.09	0	**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	

Abbreviations: FL, Florida, USVI, US Virgin Islands, PR, Puerto Rico, BLZ, Belize, PA, Panama, VEN, Venezuela, BHS, Bahamas, CUR, Curaçao, AR, Atol das Rocas, CE, Ceará, RN, Rio Grande do Norte, PE, Pernambuco, BA, Bahia, ABR, Abrolhos, ES, Espírito Santo, RJ, Rio de Janeiro, SC, Santa Catarina.

primary clades ranged from 6.58% (between *C. venezuelae* and AR *C. glaucofraenum*) to 13.29% (between Caribbean *C. glaucofraenum* and Brazil *C. glaucofraenum*; Table S2). 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. The haplotype network revealed that none of the 74 haplotypes were shared among clades and a minimum of 27 (AR *C. glaucofraenum* -*C. venezuelae*) and a maximum of 45 mutations (*C. venezuelae*-Caribbean *C. glaucofraenum* and AR *C. glaucofraenum*-*C. glaucofraenum*) connected haplotypes between clades (Figure 1). The overall star-shape configuration of the haplotype network suggests Brazil *C. glaucofraenum* and Caribbean *C. glaucofraenum* have undergone a recent expansion. As with the phylogeny above, the haplotype network showed that within *C. venezuelae*, the Venezuela population was isolated from any other Caribbean population. In contrast, Brazilian haplotypes were evenly distributed among areas with no structure detected across barriers.

3.2 | COI population genetic analyses

Basic genetic diversity estimates are summarized in Table 1. Most F_{ST} estimates among populations in different clades were significantly high (>0.91; Table 2). Conversely, populations within clades shared more gene flow as indicated by their smaller F_{ST} estimates. Populations of Caribbean *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 versus. west comparisons; t = −2.44, df = 13, p = 0.01). Populations of both *C. glaucofraenum* and *C. venezuelae* from Venezuela were highly isolated from all Caribbean populations, including strong isolation from nearby Curaçao (F_{ST} = 0.75). In contrast, most other Caribbean populations were genetically similar despite much longer distances between sites (F_{ST} = 0 – 0.20). In Brazil, the only populations that demonstrated significant levels of differentiation were Santa Catarina and Espírito Santo. Brazilian populations in the same area (north, central or south) were not significantly different from one another based on F_{ST} estimates (Wilcoxon rank-sum test; W = 50.5, p-value = 0.11).

3.3 | SNP filtering

Although 103 individuals were sent out for SNP genotyping, the final dataset included 91 samples because 12 were removed due to poor-quality sequencing. Despite failing to meet the a priori threshold for <20% missing data within an individual (Table S1), the two samples from AR were maintained in the dataset due to their importance for phylogenetic analyses. Locus outlier detection (using Bayescan v.2.1 (Foll and Gaggiotti, 2008)) revealed zero outlier loci in any of the pairwise comparisons (Figure S1), so all loci were retained for downstream analyses.

3.4 | SNP phylogenetic analyses

As with the COI tree, the Bayesian SNP tree showed strong support for four monophyletic clades (Figure 4). There was high support for the overall clade consisting of *C. venezuelae*, AR *C. glaucofraenum* and Brazil *C. glaucofraenum* 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. In contrast, samples within the Brazilian clade largely consisted of a polytomy. The maximum likelihood analysis was topologically identical at all major nodes to the Bayesian tree.

3.5 | SNP population genetic analyses

Using the ΔK approach, Bayesian clustering analyses in STRUCTURE indicated K = 2 in Caribbean *C. glaucofraenum*, although the split does not conform to any location (Figure 5). Similarly, *C. venezuelae* individuals clustered into K = 2, which also did not appear to match any known barriers. Two individuals from Curaçao (*C. venezuelae*) were strongly differentiated from the remainder of Caribbean individuals. Within Brazilian *C. glaucofraenum*, K = 2 was the most likely value, which separates the two southern populations from the remainder of Brazil and coincides with the Cabo Frio barrier. However, there was a secondary peak in likelihood that suggested K = 3 was nearly equally likely, separating Brazil into north (CE, RN and PE), central (BA, ABR and ES) and south (RJ and SC), which corresponds to the SEC and Cabo Frio barriers; no other levels of clustering were supported in Brazil. The user-generated DAPC found 20 PCs and one DF, while the cross-validated analysis had 12 PCs and one DF (Figure S2). Both DAPC results corroborated the results of STRUCTURE with Brazil having a K = 2, splitting the samples into northern-central and southern populations.

All pairwise F_{ST} estimates between populations from different clades ranged from 0.40 to 0.79 and were significant when N > 2 (Table 2). No F_{ST} estimates were significant for AR *C. glaucofraenum* because of the small sample size. However, the smallest F_{ST} comparisons between clades were found between AR *C. glaucofraenum* and Brazil *C. glaucofraenum*. Within clades, F_{ST} estimates were low (0–0.11), but many were still significantly different from zero. In Caribbean *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 region were not significantly different. In fact, populations across barriers showed significantly higher levels of F_{ST} than population pairs on the same side of a barrier (t = 4.44, df = 19, p < 0.001). Overall, Brazilian populations showed signs of limited dispersal based on the positive trend of IBD (Figure S3).

For demographic model selection using coalescent simulations, AIC identified Model 3 (Figure 2) as the most likely model (Table 3). This model suggested that an ancestral Caribbean population

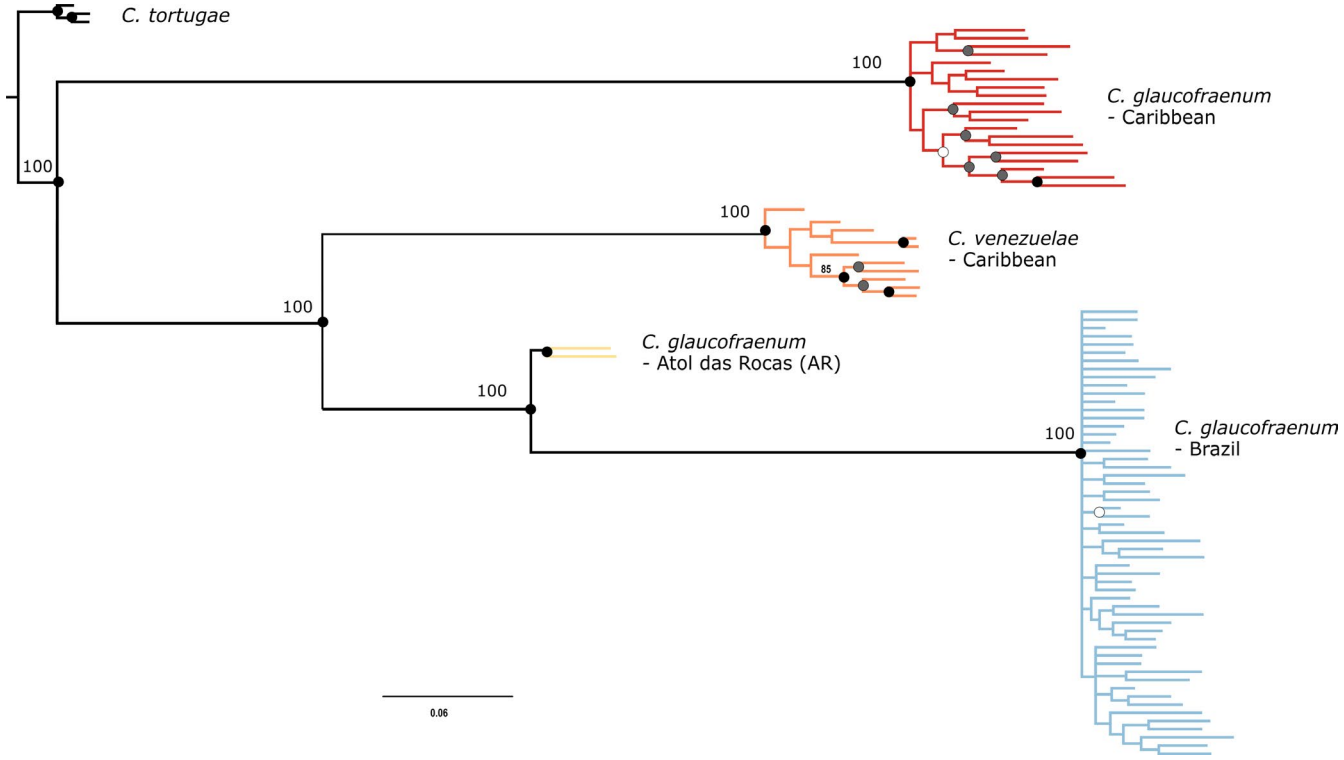

FIGURE 4 Bayesian and ML phylogeny of *C. glaucofraenum* and *C. venezuelae* using SNP data. Black, grey and white nodes represent posterior probabilities of 1.0, ≥0.95 and ≥ 0.90 respectively. Values above nodes represent bootstrap support from clades found in RAxML analysis. Each clade was constrained to monophyly 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. The ML approach used concatenated loci and a correction bias was implemented due to using all variable sites (Lewis, 2001). A GTR + G nucleotide substitution model was implemented followed by 1,000 bootstraps for likelihood estimation

dispersed to mainland Brazil. Once established along the Brazilian coast, there was an additional dispersal event to AR.

4 | DISCUSSION

In this study, we were able to demonstrate that *C. glaucofraenum* and *C. venezuelae*, exhibit significant genetic structure throughout the western Atlantic that corresponded to previously described biogeographic barriers. We used two informative datasets to find incongruence between taxonomy and evolutionary relationships. Overall, we 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. According to our coalescent simulations, a Caribbean population of *C. venezuelae* likely colonized Brazil first, followed by dispersal from Brazil to AR. In addition, several population-level barriers were found including an east-west Caribbean divide, isolation of Venezuela from the rest of the Caribbean, the cold-water upwelling at Cabo Frio and possibly the weak barrier from the southern equatorial current (SEC). These results are discussed in more detail below as they relate to phylogeography of marine taxa in the western Atlantic.

4.1 | Phylogeny and taxonomy

Even though each monophyletic clade was strongly supported, the relationships among these lineages were discordant between the SNP and COI data. The mtDNA suggests AR *C. glaucofraenum* and *C. venezuelae* are more closely related, but SNP data suggests AR *C. glaucofraenum* and Brazil *C. glaucofraenum* 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 et al., 2005; Pinheiro et al., 2018; Rocha et al., 2005), the relationships determined using SNPs more likely represent an accurate representation of the species tree for two reasons. First, the proximity of AR and the 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) is likely to infer a more accurate species tree overall and resolve homoplasy caused by either introgression or incomplete lineage sorting that is likely to occur when analysing only a single (mitochondrial) gene (Brito & Edwards, 2009).

We found the closely related Brazilian and AR lineages were likely formed as a result of dispersal and subsequent isolation from the Amazon River outflow rather than a vicariant event where the Amazon River restricted gene flow between a continuous

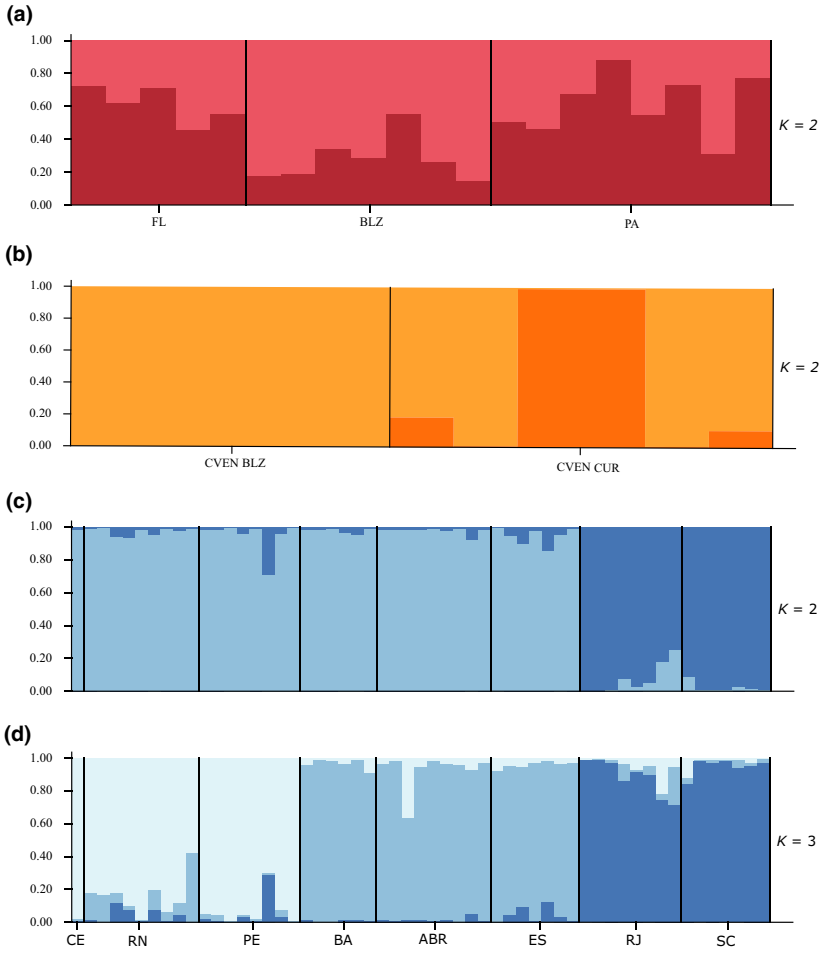

FIGURE 5 Results from STRUCTURE for (a) Caribbean *C. glaucofraenum* (FL, BLZ, PA), (b) *C. venezeuelae* (CVEN BLZ and CVEN CUR) and Brazil *C. glaucofraenum* (CE, RN, PE, BA, AB, ES, RJ, SC) for (c) K = 2 and (d) K = 3

TABLE 3 Summary of model selection results. Model 3 has the highest likelihood and lowest Δ_i making it the best model of the four tested

Model	Max $\log_{10}(Lhood_i)$	No. of parameters	AIC_i	Δ_i
1	−1714.389	5	3,438.778	1,189.568
2	−1295.332	6	2,602.664	353.454
3	−1119.605	5	2,249.21	0
4	−1136.871	5	2,283.742	34.532

population. The Amazon River is a well-known ecological barrier for many marine taxa and often results in speciation for low dispersal organisms like *C. glaucofraenum* and *C. venezuelae* (Briggs & Bowen, 2012; Dias et al., 2019; Floeter et al., 2008). The Amazon River likely intermittently restricted gene flow for the past 9 Myr due to fluctuating sea levels that created or prevented dispersal corridors (Hoorn et al., 2017; Rocha, 2003). Based on the permeable nature of marine barriers, opportunities likely existed for intermittent dispersal across the Amazon, followed by periods of minimal gene flow that could result in speciation. Given that Brazil and AR diverged after Caribbean *C. glaucofraenum* and *C. venezuelae* split approximately 4.21 Myr (Tornabene et al., 2013), this would suggest that the Amazon barrier could have impacted divergence between Brazil-AR and *C. venezuelae*. This begs the question, why is AR so differentiated from the mainland? Several studies have suggested that endemism is caused by ecological differences between coastal and oceanic populations (e.g. 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 here. Additional morphological and ecological data are needed to more thoroughly differentiate these taxa.

4.2 | Caribbean biogeography

There are two synergistic hypotheses to explain why marine biodiversity is high in the Caribbean: 1) the Caribbean serves as a centre of origin for marine speciation in the western Atlantic (Floeter et al., 2008) and 2) the Caribbean serves as a centre of accumulation from nearby areas (Bowen et al. 2013; Rocha et al. 2008). In the present study, we can deduce that dispersal occurred out of the Caribbean to Brazil. The majority of *Coryphopterus* species occur in the Caribbean and only three species exist outside the Caribbean. Indeed, the fact that *C. venezuelae* occur at a depth of 69m below sea level (Baldwin & Robertson, 2015) suggests a possible mechanism for dispersing beyond the Amazon River and inhabiting the

extensive mesophotic reef systems at the mouth of the Amazon River (Francini-Filho et al. 2018).

Within-clade analyses showed distinct patterns of barriers impacting connectivity throughout 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 among taxa (DeBiasse et al., 2016; Foster et al., 2012; Taylor & Hellberg, 2006). Regarding the present study, we found evidence for a similar division among populations of *C. glaucofraenum* in the Caribbean. While this pattern was not clear for *C. venezuelae*, increased sampling may yield more statistical power to test this barrier. The clear demarcation of Venezuela from other Caribbean populations for both *C. glaucofraenum* and *C. venezuelae*, including a nearby population in Curaçao (*C. venezuelae*), has not been observed previously. However, similar isolation across short distances has been found in other parts of the Caribbean due to local currents (Foster et al., 2012; Jackson et al., 2014). Strong currents and the thin continental shelf near Venezuela may lead to larvae that are lost offshore resulting in low connectivity for both species near Venezuela (e.g. D'Agostini et al., 2015).

4.3 | Brazilian biogeography

In accordance with our prediction, two possible 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 been referenced as a possible barrier for *Millepora* fire corals and for *Symbiodinium* dinoflagellates in the scleractinian coral *Mussismilia hipida* (Souza et al., 2017; Picciani et al., 2016). Second, the genetic break occurs between 8° and 13°S and could be due to current bifurcations. D'Agostini et al. (2015) found that larvae flowed in opposite directions in different seasons due to variability in the SEC. Although there is no information on *Coryphopterus* spawning in Brazil, multiple spawning events of *C. glaucofraenum* can occur over an extended timeframe which likely overlap with the change in the SEC (Forrester et al., 2010).

The more evident barrier was 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 many taxa (Spalding et al., 2007) and has been known to cause differentiation in crustaceans and fishes (Boschi, 2000; Machado et al., 2017; Santos et al., 2006). Two possible reasons for differentiation across the Cabo Frio barrier are 1) ecological differences across the barrier and 2) currents that prevent larval dispersal. First, the cold water and nutrient upwelling system represents an ecological transition from tropical coral reefs northward to warm temperate rocky reefs southward (Ferreira et al., 2004; Santos et al., 2006). Marceniuk et al. (2019) suggest this ecological transition led to hybridization after secondary contact between *Orthopristis ruber* and *O. scapularis*. Ecological transitions were also suggested to have caused tropical and subtropical clades to diverge in *Chaetodipterus faber* (Machado et al., 2017). Second, ocean currents may physically restrict gene flow between central and southern populations. There is a tendency for the Brazil Current to lose pelagic larvae off the continental shelf rather than follow the coastline and maintain connectivity between central and southern Brazil (D'Agostini et al., 2015; Endo et al., 2019). The similarity in biogeographic patterns across many species along Brazil is intriguing and warrants further investigation among different taxa.

Overall, this study has demonstrated how the biogeography of a western Atlantic goby species complex was influenced by genetic connectivity through permeable biogeographic barriers. The Amazon River outflow has isolated Brazilian from Caribbean lineages while the Brazilian oceanic reef lineage has also diverged from the coastal lineage. Furthermore, both COI and SNP datasets provided important information regarding 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/São Francisco and Cabo Frio barriers in Brazil that were undetectable using a single coarse marker. Lastly, coalescent simulations shed light on the processes that led to observed phylogeographic patterns. This study provides groundwork for future seascape genetics studies to evaluate how different oceanographic features impact geneflow among populations. We suggest additional studies be performed in the western Atlantic across taxa to confirm the processes at work with particular emphasis along Brazil's coast.

5 | Author Biosketch

DRV conducted this project as a MS student in the laboratory of EAH. CELF works with many aspects of marine biology to understand how reef systems function, including comparative issues of Atlantic biogeography and global macroecology. SRF works with biogeography, macroecology and evolution of reef fishes. JDK is a PhD student in the laboratory of EAH with a focus on evolutionary adaptation to novel environments. EAH's laboratory uses molecular genetic techniques to answers questions associated with genetic structure and species boundaries. As a team, all authors work to integrate how ecological and evolutionary implications impact species boundaries and connectivity among populations of Atlantic fishes. This team combines the expertise of CELF and SRF with regard to the evolutionary ecology of fishes with the genetic and modelling approaches of DRV, EAH and JDK.

ACKNOWLEDGEMENTS

We would like to thank Ramon Noguchi, Ben Victor, Raphael Macieira and Carole Baldwin for providing samples for this study. R. Macieira helped with species identification. We would like to thank

Michelle Gaither and two anonymous reviewers for helpful edits on the manuscript. This project was supported by the UCF Department of Biology and the Brazilian Marine Biodiversity Network – SISBIOTA-Mar (P.I: S.R. Floeter; CNPq 563276/2010-0 and FAPESC 6308/2011-8) under the following permit: SISBIO 29953-1.

AUTHOR CONTRIBUTIONS

DRV performed laboratory work, analysed data and wrote the manuscript. JDK performed FASTSIMCOAL2 analyses and helped edit the manuscript. DRV, SRF, CELF and EAH conceived the idea, provided samples and edited the manuscript.

DATA AVAILABILITY STATEMENT

COI sequence data have been submitted to the GenBank database under accession number MT627483–MT627594. SNP data are archived at https://datadryad.org/stash/share/qt9DrBZIRGIrRl_fpFPvq-NYeNGaX12X4bhNMap_Xxk

ORCID

Eric A. Hoffman https://orcid.org/0000-0002-2432-3619

REFERENCES

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. https://doi.org/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* (pp. 111–138).

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.

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., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*. https://doi.org/10.1371/journal.pcbi.1003537

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

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

D'Agostini, A., Gherardi, D. F. M., & Pezzi, L. P. (2015). Connectivity of marine protected areas and ts relation with total kinetic energy. *PLoS One*, *10*, e0139601.

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

de Souza, J. N., Nunes, F. L. D., Zilberberg, C., Sanchez, J. A., Migotto, A. E., Hoeksema, B. W., Serrano, X. M., Baker, A. C., & Lindner, A. (2017). Contrasting patterns of connectivity among endemic and widespread fire coral species (*Millepora* spp.) in the tropical Southwestern Atlantic. *Coral Reefs*, *36*, 701–716. https://doi.org/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*, 2136–2146. https://doi.org/10.1111/jbi.12785

Dias, R. M., Lima, S. M. Q., Mendes, L. F., Almeida, D. F., Paiva, P. C., & Britto, M. R. (2019). Different speciation processes in a cryptobenthic reef fish from the Western Tropical Atlantic. *Hydrobiologia*, *837*, 133–147. https://doi.org/10.1007/s10750-019-3966-z

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*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Endo C. A. K., Gherardi D. F. M., Pezzi L. P., & Lima L. N. (2019). Low connectivity compromises the conservation of reef fishes by marine protected areas in the tropical South Atlantic. *Scientific Reports*, *9*, 8634. https://doi.org/10.1038/s41598-019-45042-0

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

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust Demographic Inference from Genomic and SNP Data. *PLoS Genetics*, *9*. https://doi.org/10.1371/journal.pgen.1003905

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*, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

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*, 1093–1106. https://doi.org/10.1111/j.1365-2699.2004.01044.x

Floeter, S. R., Rocha, L. A., Robertson, D. R., Joyeux, J. C., Smith-Vaniz, W. F., Wirtz, P., Edwards, A. J., Barreiros, J. P., Ferreira, C. E. L., Gasparini, J. L., Brito, A., Falcon, J. M., Bowen, B. W., & Bernardi, G. (2008). Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, *35*, 22–47. https://doi.org/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*, 155–163. https://doi.org/10.1007/s10144-010-0225-6

Foster, N. L., Paris, C. B., Kool, J. T., Baums, I. B., Stevens, J. R., Sanchez, J. A., Bastidas, C., Agudelo, C., Bush, P., Day, O., Ferrari, R., Gonzalez, P., Gore, S., Guppy, R., McCARTNEY, M. A., McCOY, C., Mendes, J., Srinivasan, A., Steiner, S., … Mumby, P. J. (2012). Connectivity of Caribbean coral populations: Complementary insights from empirical and modelled gene flow. *Molecular Ecology*, *21*, 1143–1157. https://doi.org/10.1111/j.1365-294X.2012.05455.x

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

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*, *153*, 51–65. https://doi.org/10.1016/j.gloplacha.2017.02.005.

Huyghe, F., & Kochzius, M. (2018). Sea surface currents and geographic isolation shape the genetic population structure of a coral reef fish in the Indian Ocean. *PLoS One*, *13*(3), e0193825.

Jackson, A. M., Semmens, B. X., Sadovy de Mitcheson, Y., Nemeth, R. S., Heppell, S. A., Bush, P. G., Aguilar-Perera, A., Claydon, J. A. B., Calosso, M. C., Sealey, K. S., Schärer, M. T., & Bernardi, G. (2014).

Population Structure and Phylogeography in Nassau Grouper (*Epinephelus striatus*), a Mass-Aggregating Marine Fish. *PLoS One*, *9*. https://doi.org/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*, *33*, 1870–1874. https://doi.org/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*, 1695–1701. https://doi.org/10.1093/molbev/mss020

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

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

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*, 211–223. https://doi.org/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*, 298–299. https://doi.org/10.1093/bioinformatics/btr642

Machado, L. F., Damasceno, J. 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. https://doi.org/10.1016/j.jembe.2016.11.005

Marceniuk, A. P., Caires, R. A., Machado, L., Cerqueira, N. N. C. D., Serra, R. R. M. S., & Oliveira, C. (2019). Redescription of *Orthopristis ruber* and *Orthopristis scapularis* (Haemulidae: Perciformes), with a hybridization zone off the Atlantic coast of South Amerca. *Zootaxa*, *4576*, 109–126.

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*, 5514–5523. https://doi.org/10.1002/ece3.3161

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*. https://doi.org/10.1007/s00227-016-3010-z

Pinheiro, H. T., Rocha, L. A., Macieira, R. M., Carvalho-Filho, A., Anderson, A. B., Bender, M. G., Di Dario, F., Ferreira, C. E. L., Figueiredo-Filho, J., Francini-Filho, R., Gasparini, J. L., Joyeux, J.-C., Luiz, O. J., Mincarone, M. M., Moura, R. L., Nunes, J. D. A. C. C., Quimbayo, J. P., Rosa, R. S., Sampaio, C. L. S., Sazima, I., Simon, T., Vila-Nova, D. A., & Floeter, S. R. (2018). South-western Atlantic reef fishes: Zoogeographical patterns and ecological drivers reveal a secondary biodiversity centre in the Atlantic Ocean. *Diversity and Distributions*, *24*, 951–965. https://doi.org/10.1111/ddi.12729

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Rocha, L. A. (2003). Patterns of distribution and processes of speciation in Brazilian reef fishes. *Journal of Biogeography*, *30*, 1161–1171. https://doi.org/10.1046/j.1365-2699.2003.00900.x

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*, 573–579. https://doi.org/10.1098/2004.3005

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

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*, 1572–1574. https://doi.org/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*, 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

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. https://doi.org/10.7717/peerj.1106

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*, 4361–4373. https://doi.org/10.1111/j.1365-294X.2006.03108.x

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A. L., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience*, *57*, 573–583. https://doi.org/10.1641/B570707

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

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*, 695–707. https://doi.org/10.1111/j.1365-294X.2006.02820.x

Tigano, A., & Friesen, V. L. (2016). Genomics of local adapation with gene flow. *Molecular Ecology*, *25*, 2144–2164.

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

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*, 1847–1857. https://doi.org/10.1098/rstb.2005.1716

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Volk DR, Konvalina JD, Floeter SR, Ferreira CEL, Hoffman EA. Going against the flow: Barriers to gene flow impact patterns of connectivity in cryptic coral reef gobies throughout the western Atlantic. *J Biogeogr.* 2020;00:1–13. https://doi.org/10.1111/jbi.14010